NEH Application Cover Sheet
Digital Humanities Start-up Grants

PROJECT DIRECTOR

Dr. Harley Hamilton
Senior Research Scientist
85 5th Street
Atlanta, GA 30308-0204
UNITED STATES

E-mail: harley.hamilton@cc.gatech.edu
Phone(W): 404-385-1570
Phone(H):
Fax:

Field of Expertise: Technical: Education

INSTITUTION

Georgia Tech Research Corporation
Atlanta, GA UNITED STATES

APPLICATION INFORMATION

Title: *Dictionary 3.0*

Grant Period: From 5/2014 to 10/2014
Field of Project: Languages: English

Description of Project: Monolingual multimedia English dictioanry

BUDGET

Outright Request	$59,990.00	**Cost Sharing**	$0.00
Matching Request	$0.00	**Total Budget**	$59,990.00
Total NEH	$59,990.00		

GRANT ADMINISTRATOR

Mrs. Stacey Oliver-Gooden
Contracting Officer
Office of Sponsored Programs
505 Tenth Street
Atlanta, GA 30332-0420
UNITED STATES

E-mail: sogooden@gatech.edu
Phone(W): 404-894-6930
Fax: 404-894-5945

1. **Table of contents**

2. **List of participants**
Harley Hamilton- Georgia Institute of Technology
Scott Robertson- Georgia Institute of Technology
Don McCreary- University of Georgia
Daren Brabham- University of Southern California

3. **Abstract**

Dictionary 3.0 (D3) will be a free multimedia English language dictionary that provides a powerful reference tool for students, adult second language learners, and life-long learners, essentially anyone. As a humanities reference tool it will be an invaluable living repository of the English lexicon as it will be continually grow and be updated via crowdsourcing. D3 will revolutionize the presentation format of dictionaries which has changed little since Samuel Johnson's *A Dictionary of the English Language* (1755). Even today, text dominates as the presentation medium, yet technology allows much more. D3's multimedia presentation will include not only text to represent meaning but text-to-speech capabilities, audio, and video to illustrate word meaning in a rich context, just as word meaning is embodied in everyday language usage. <u>Meaning</u> is what dictionary users most desire (Lew 2010). D3 will deliver word meaning that is accessible on any device, computer, tablet, phone…
<div align="center">"Any word, anywhere, anytime"</div>

Statement of Innovation

New approaches to dictionary-making practice are needed as the field of lexicography seeks to invigorate itself (Alonso, Millon, & Williams, 2011). Dictionary design today should be dominated by "innovation and bold intelligence" not tradition (Gouws (2011). A dictionary should be a monofunctional tool focused solely on providing word meaning (Leroyer, 2011; Berghenholtz & Bergenholtz, 2011). D3 will be such an innovative and boldly intelligent tool.

Statement of Humanities Significance

The humanities include dictionary creation, language study, and history of the arts. D3 will provide context rich examples of the English lexicon via video clips that may serve linguistic study of usage. As clips will be culled from video produced over the last 100 years it will also provide scholars with a historical database of language usage and preserve examples of the visual and performing arts over that time. Crowdsourcing will allow for everyman to contribute to this humanities reference.

4. **Narrative**

o **Enhancing the humanities through innovation**

We are seeking a Phase I, Level II grant to plan and develop a prototype of D3. This work will lead to a functional set of apps and tutorials that will be a complete infrastructure for the fulfillment of D3 in the next phase grant which will include creation of the D3 lexicon by grant staff and the public via crowdsourcing. The creation of D3 will revolutionize the field of dictionary making while providing an educational tool to assist students with language learning and reading. It will also provide a participatory humanities activity (crowdsourcing) that allows anyone to contribute to the creation of a humanities reference work containing the lexicon of English. The creation of D3 entries by the general public can be accomplished by open-source/freely accessible tools. People will be able to create contributions in several ways,

1. Using a digital camera, camcorder, or phone to capture video and post it to YouTube. These devices are common and owned by a vast majority of individuals.
   Their video can be edited in several free editing programs such as the YouTube editor, Windows Movie Maker, and iMovie as well as other free options
2. Using Creative Commons media which are integrated with the free YouTube editing app to create new videos. There are over 4 million videos and hundreds of millions of still images available with Creative Commons licensing that allows for re-use.
3. Free tools such as Muvizu, FreshBrain, and WeGame are available for making animated videos.

Workflow

Contributor tasks

- Establish own YouTube channel.
- Create entry and upload to YouTube channel
- Submit entry to D3 via web form (see Appendix A)

D3 editorial tasks- performed initially by project director for user studies until volunteer editorial network is developed as part of later grant development and funding.

- Review submitted entry to accept, reject, or comment for revision and resubmission
  o If accepted, (see Appendix B for acceptance criteria)
     ▪ Accepted video is uploaded to D3's YouTube channel
     ▪ The word, definition, sample sentence, and video url are entered in the D3 database for access via D3 user app.
  o If rejected, contributor is informed of the D3 decision and the reason for rejection
  o If revisions are suggested the entry will then be reconsidered upon resubmission

Crowdsourcing is an economical and common method for dictionary development. The sheer magnitude of a project capable of producing a dictionary covering the words of a language precludes the possibility of any single publisher creating such a tool due to time and cost-restraints (Fillmore, 2003). This is even more pertinent when all entries contain audio and video, not simply text. Scaling such a tool requires a crowd (Weld et al. 2012) and has been successful in the past. In 1859, the British Philological Society launched an appeal to the British and American public to create a new dictionary "worthy of the English Language and of the present state of philological science." Over 6 million submissions were made over a period of 70 years (Winchester, 1998). This helped produce the Oxford English Dictionary. Today, the Oxford English dictionary is still employing this tactic. Other dictionary makers also employ crowdsourcing

http://www.macmillandictionary.com/us/open-dictionary/

http://nws.merriam-webster.com/opendictionary/submit.php

http://www.ahdictionary.com/word/submitNew.html

D3 follows in their footsteps to create a dictionary worthy of the English language and the present state of technological as well as philological science. To facilitate crowd-sourcing, we will rely on both intrinsic and extrinsic motivations. These include fun, learning, recognition, self-satisfaction from contributing to a large project of common interest, altruism, and creative expression (Borst, 2010; Brabham, 2012). D3 is a hybrid Wiktionary and YouTube entity, both of which are populated by contributions by users enticed by the motivations just mentioned. Contributing to YouTube is a rather easy task as evidenced by the sheer volume of video contributed. One hundred hours of video are uploaded to YouTube each minute.

o **Environmental scan**

Dictionaries of today
Dictionary 1.0 (D1)
Before the technological revolution and birth of the internet, dictionaries were massive text-based paper tomes. These still exist today just as the coelacanth remains extant (see The American Heritage Student Dictionary, 2012). "Illustrated" dictionaries include a few images but remain text-dominant (see Merriam-Webster Elementary Dictionary, 2010). "Picture" dictionaries for younger students are available but the content is generally limited to easily imaged content such as nouns, action verbs, and some adjectives (see Scholastic First Picture Dictionary, 2009). Words such as *spooky* (found in a 1st grade reader, Afflerbach, et al., 2000) would be difficult if not impossible to clearly illustrate with a static image and are not included. However, the word *blue tit* (a bird) is. Each of the aforementioned dictionaries is rated as 5 out of 5 stars on Amazon.com.
Finding words in paper dictionaries is also problematic and time consuming as a fairly high level of alphabetizing skills is needed. Studying children age 7-11, Beech (2004) found that for younger readers (age 7-9.4) the average look up time for a printed word was 48 seconds with an accuracy rate of 68%. For older children (age 9.5-11) the average look up time was 28 seconds with an average accuracy rate of 90%. Clearly, for many children, finding words in a paper dictionary is time consuming and an inefficient, inaccurate means of accessing word meaning. The search time plus encountering unknown words in the definitions themselves take attention and memory resources away from the reading passage for a rather long period of time and may overload working memory and negatively impact reading comprehension (Denh, 2008; Alloway, Gathercole, Kirkwood, & Elliott, 2009).
Dictionary 2.0 (D2)
D2 or electronic dictionaries may be web-based, on CD/DVDs, or stand-alone devices. The limitations of the early internet (Web 1.0) restricted D2 to mainly text-based entries. CD/DVDs allowed for multimedia entries but were still limited due to the large size of video and image files and the relatively small storage size of the disks themselves. CD/DVDs are also limited in ease of accessibility due to the need for a computer to utilize them. These media are also not cost effective as they range in price from $10 to over $200 . Stand-alone devices such as

those made by Franklin range in price from $50 to $400. They are predominantly text-based due to technology limitations and cost.

As technology developed and the internet evolved, D2 began to appear online. As Web1.0 became Web 2.0 with full multimedia capabilities dictionary development, however, has remained stagnant. An online dictionary may be just a copy of the paper dictionary (see www.webster-dictionary.org/). The top two online adult dictionaries as determined by a Google search (see www.merriam-webster.com/, and www.dictionary.com) do include a minor multimedia component that allows the user to hear only the target word spoken. The user interface of each of these is extremely cluttered with ads and other features such as links to quizzes, word of the day, a thesaurus, Spanish…. These two sites also include sample sentences. However, the sample sentences are inadequate unless you already know the meaning of the target word. For example at www.dictionary.com a sample sentence for *spooky* is "They make trees look spooky, and they prefer not to be disturbed." The enigmatic nature of this sentence is overwhelming, perhaps even spooky.

Visual dictionaries have appeared, yet are limited to static images, primarily of nouns. The breadth of the largest online visual dictionary (http://visual.merriam-webster.com/ is 20,000 words and 6,000 images. *Spooky* cannot be found. This may cause users frustration due to lack of a substantial lexicon. Visual webs such as http://www.visualthesaurus.com/ are also available yet remain exclusively text-based.

Today, most web-based dictionaries remain text-based or include only minor multimedia components. While access time and accuracy of search may be improved by allowing the user to type the target word in a text box, current on-line dictionaries do not take full advantage of the potential offered by Web 2.0 technology (deSchryver, 2003). Such a statement was true a decade ago and remains accurate to this day.

Dictionary 2.5 (D2.5)

Some attempts are being made to utilize the multimedia power of Web 2.0. Wordia primarily utilizes video of people reading the definition or explaining the word alongside its text definition. Meaning is primarily rendered by the spoken word. It is crowdsourced and requests users to upload entries when video related to a text definition does not exist. Its editorial policy appears to allow for esoteric tributes rather than clear representations of meaning. See a definition for *dog* contributed by nickgrim (lower left hand corner of web page).

Vidtionary contains video and is limited to easily represented nouns and verbs (although the word *cat* could not be found). Their video definitions can be confusing, search for *color* to see an example. Once again *spooky* is unavailable in each, even though the media lends itself to a vivid depiction of the meaning of *spooky*.

Dictionary 3.0 (D3)

D3 will take advantage of the multimedia capabilities of Web 2.0, the popularity of inexpensive devices such as phones, tablets, and laptops, and the fully connected nature of today's technological infrastructure to become an inexpensive, media-rich, ubiquitous reference work on the English lexicon that brings word meaning to those who need it "any word, anywhere, anytime". Word meaning will be illustrated via video rich in physical and social context clearly depict meaning, just as in everyday life. See *spooky* . The meaning of *spooky* in this example is illustrated by both the video and its accompanying spooky audio as well as the definition and sample sentence which are directly related to the video.

o **History and duration of the project**

D3 began in January, 2013. It is the evolution of the project director's work on the English to American Sign Language bilingual dictionary, SMARTSign. Work on the SMARTSign dictionary was awarded the American Library Association's Grolier Foundation award for making the greatest contribution to children's literacy in 2004. A review of the literature for a study of the efficiency of dictionary use by deaf signing students (Hamilton, 2012) made the project director aware of the impact of vocabulary on reading for all readers, hearing and deaf. This review also made it clear that while dictionaries are powerful learning tools they are sorely underutilized due to their current format. That information along with knowledge of what current technology can provide melded into the idea of D3. When demos have been done for a wide range of consumers, the general response has been "That is so cool!" (even by people my age, 60). Enlisting collaborators from academia has been easy and elicited the same response.

D3 currently has no financial support. As a research scientist at Georgia Tech, the project director is seeking funding for this project from NEH and other governmental sources. We will continue to seek funding to make D3 a living, growing, social media platform that provides an avenue for participants to collectively create a complete reference work of the English lexicon as well as a powerful learning tool for students.

o **Work plan**
This work will include development of
- D3 back-end database. Design back-end database schema and data models using an open-source database such as MongoDB or MySQL. The database schema will include fields for storing dictionary terms, their associated URL links, text definition and sample sentences, user interaction logging including time-stamped entries per user, dictionary terms searched for, and subsequent hits and misses. Person responsible- Scott Robertson
- Server API design: Design and implement a back-end web service application programming interface (API) using a REST and JSON software architecture. This API will allow client applications (apps) to create, read, update and delete dictionary database records as well as create and read user interaction log records. This API will support an administrative web app, used to administer the dictionary database (i.e. add new entries), client web and mobile apps and desktop browser widgets or extensions. Person responsible- Scott Robertson
- Client mobile app design: Design and implement a cross-platform, HTML5 and Javascript/AJAX based mobile app and user interface to allow users to search for dictionary terms and view the targeted entry. This app will be developed using a development framework such as Sencha Touch in order to run on multiple platforms: iOS, Android, Blackberry, Windows Mobile. Such HTML5 apps can be packaged and delivered as native device apps or can be run in a device's mobile web browser. Person responsible- Scott Robertson
- Desktop browser widget and/or extension: Design and develop a web browser extension and user interface which will implement the following functionality: Allow a user to highlight any word on a web page she is browsing and perform a D3 dictionary search, presenting the search result video(s) in a pop-up browser window. This extension will initially be developed for the Google Chrome browser using the Google Chrome

extensions API. Google Chrome is a cross-platform browser and Chrome extensions can be used on Windows and Apple OSX platforms. Person responsible- Scott Robertson

- Crowdsourcing app- A web-based form will be developed using HTML and javascript that will allow people to submit entries to D3 for acceptance or editorial suggestions. This will include word, definition, sample sentence, video url, and contact information. Contributors will store their videos on their own YouTube channel. If an entry is accepted it will then become part of D3 and stored on the D3 YouTube channel to ensure availability. Person responsible- Scott Robertson
- D3 lexicon- Identify scope of D3 lexicon and provide a web-based resources for suggested entries for crowdsourcing participants. Persons responsible- Don McCreary, Harley Hamilton
- Sample D3 entries- Five hundred D3 entries will be created. Person responsible- Harley Hamilton
- Video tutorials- YouTube-based tutorials showing various ways to create and submit D3 entries using several different free tools such as Windows Movie Maker, iMovie, Muvizu, WeGame, YouTube video editor, Creative Commons content, and video camera. Person responsible- Harley Hamilton
- Crowdsourcing mechanism- Develop means to inform the public of D3 and encourage participation in its creation as well as the means for submitting entries. Person responsible- Daren Brabham, Harley Hamilton
- Update D3 webpage- As D3 develops, its webpage will be used as a major source of new information for social and mainstream media outlets. Person responsible- Harley Hamilton
- User studies- Studies to iteratively refine the interface, crowdsourcing mechanism, and tutorials. These will inform the development of D3. See attached studies. IRB clearance will be obtained via normal Georgia Tech channels. This entails completion of a detailed IRB request made online to the Office of Research Integrity Assurance http://researchintegrity.gatech.edu/. Letters of support from local school districts to participate in the studies are in appendix C. Person responsible- Harley Hamilton

Sustainability
- Technology- All apps will reside and be backed up on secure servers at Georgia Tech. The videos will be hosted by YouTube via the D3 YouTube channel. Backup copies of all video will be stored on secure Georgia Tech servers.
- D3 editing and updating- If the initial grant is funded, a subsequent full grant will be applied for. During the second grant period, funds will be requested for two research assistants whose responsibilities will include the editorial task for accepting entries and including these in the D3 lexicon. If an entry passes both assistants, the project director will have the final say on an entry's acceptance. During this second grant, a crowdsourced editorial board will be assembled following the lead of Wikipedia and Wiktionary to ensure continued D3 growth after funding is no longer available.

Staff

Harley Hamilton, (Project director)- Dr. Hamilton will oversee the project as well as create sample entries and tutorials, and conduct user studies to iteratively refine D3. He will work on D3 1.5 months over the one year grant period.

Scott Robertson, (Lead developer)- Mr. Robertson will develop all apps for D3 as described above. He will devote two months to D3 over the course of the grant.

Don McCreary, (Dictionary consultant)- With Dr. Hamilton, Dr. McCreary will formulate the initial lexicon to be crowdsourced in order to populate D3.

Daren Brabham, (Crowdsourcing consultant)- Dr. Brabham will design the plan and mechanisms for implementing D3 crowdsourcing.

o **Final product and dissemination**

Social media will be used to generate "buzz" so that future crowdsourcing can hit the ground running. D3 already has Twitter, YouTube, and Facebook accounts established and these will be utilized to promote it. Academic articles describing the development of D3 will be submitted to the International Journal of Lexicography and Lexicographica. Press releases will be sent to all major news outlets via the Georgia Tech media relations department. The D3 webpage will be updated and used to refer people to via tweets and press releases

Title: Dictioanry 3.0									
Estimated Start Date:	1/1/2014								
Sponsor:	GT starter								
	Year 1	Year 2	Year 3	Total		MONTHS			
PERSONAL SERVICES					Salary	Yr 1	Yr 2	Yr 3	
Faculty	$14,000	$0	$0	$14,000	7,000	2			Robertson-programming
Faculty	$0	$0	$0	$0					
Faculty	$0	$0	$0	$0					
Faculty	$0	$0	$0	$0					
Total Senior Personnel	$14,000	$0	$0	$14,000					
Post Doc	$0	$0	$0	$0	4,583				
Research Scientist	$0	$0	$0	$0	5,000				
TBA-mo	$0	$0	$0	$0	4,341				
Research scientist	$10,500	$0	$0	$10,500	7,000	1.5			Hamilton-director
GRAs @ 45%	$0	$0	$0	$0	2,131				
GRA @ 33%	$0	$0	$0	$0	1,579				
Secretarial	$0	$0	$0	$0	3,000				
Total Other Personnel	$10,500	$0	$0	$10,500					
Total Personal Services	$24,500	$0	$0	$24,500					
FRINGE									
@27.9%	$3,906	$0	$0	$3,906					
@0.018.%	$189	$0	$0	$189					
Total P.S. & Fringe	$28,595	$0	$0	$28,595					
EQUIPMENT				$0					
TRAVEL	$500			$500					NEH meeting
OTHER DIRECT COSTS									
Materials & Supplies	$2,000			$2,000					software
Publication Costs				$0					
GRA Tuition $1170/mo	$0	$0	$0	$0					
Consulting	$7,000	$0	$0	$7,000					Brabham, $5000; McCreary, $2000
Computing Charges	$1,191	$0	$0	$1,191					
Total Other Direct	$10,191	$0	$0	$10,191					
TOTAL DIRECT COSTS	$39,286	$0	$0	$39,286					
INDIRECT									
@ 52.7	$20,704	$0	$0	$20,704					
TOTAL DIRECT & INDIREC	$59,990	$0	$0	$59,990					

DEPARTMENT OF THE NAVY
OFFICE OF NAVAL RESEARCH
875 NORTH RANDOLPH STREET
SUITE 1425
ARLINGTON, VA 22203-1995

IN REPLY REFER TO

Agreement Date: June 25, 2013

NEGOTIATION AGREEMENT

INSTITUTION: **GEORGIA INSTITUTE OF TECHNOLOGY**
GEORGIA TECH RESEARCH CORPORATION
ATLANTA, GEORGIA 30332

The fringe benefit rates contained herein are for use on grants, contracts and/or other agreements issued or awarded to the Georgia Institute of Technology/Georgia Tech Research Corporation (GIT/GTRC) by all Federal Agencies of the United States of America, in accordance with the provisions and cost principles mandated by 2 CFR 220 (formerly OMB Circular A-21). These rates shall be used for forward pricing and billing purposes for GIT/GTRC's Fiscal Year 2014. This rate agreement supersedes all previous rate agreements for Fiscal Year 2014.

Section I: RATES - TYPE: PROVISIONAL (PROV)

Fringe Benefits Rates:

CATEGORY	TYPE	FROM	TO	RATE	BASE	APPLICABLE TO	LOCATION
Full Benefits	Prov	7/1/2013	6/30/2014	28.5%	(a)	All Programs	RI*
Limited Benefits	Prov	7/1/2013	6/30/2014	19.5%	(b)	All Programs	RI*
Partial Benefits	Prov	7/1/2013	6/30/2014	1.4%	(c)	All Programs	RI*
Grad Student Health	Prov	7/1/2013	6/30/2014	1.9%	(d)	All Programs	RI*

*Resident Instruction

DISTRIBUTION BASES

(a) Salaries and wages of (i) regular full-time faculty, (ii) principal investigators, (iii) professional and administrative staff, (iv) joint staff, (v) temporary academic or research professionals participating in Institute retirement programs and group health and life insurance, (vi) bi-weekly permanent employees, and (vii) part-time employees who work at least 75 percent but less than 100 percent of a full-time work schedule.

(b) Salaries and wages of part-time employees who work at least 50 percent but less than 75 percent of a full-time work schedule.

(c) Salaries and wages of employees who participate in all or part of the social security program but do not participate in retirement or group health and life insurance plans. This rate covers (i) temporary classified persons, (ii) temporary academic or research professionals not eligible for the retirement programs or group health or life insurance coverage, (iii) student employees who are registered for less than a full academic load, and (iv) part-time employees employed for less than 50 percent of a full work schedule.

(d) Salaries and wages of Graduate Research and Graduate Teaching Assistants who qualify for the graduate tuition remission award and who provide services to GIT at 33% to 50% time while being registered for at least 12 credit hours.

SECTION II - GENERAL TERMS AND CONDITIONS

A. **LIMITATIONS:** Use of the rates set forth under Section I is subject to any statutory or administrative limitations and is applicable to a given grant, contract or other agreement only to the extent that funds are available and consistent with any and all limitations of cost clauses or provisions, if any, contained therein. Acceptance of any or all of the rates agreed to herein is predicated upon all the following conditions: (1) that no costs other than those incurred by the grantee/contractor were included in its indirect cost pool as finally accepted and that all such costs are legal obligations of the grantee/contractor and allowable under governing cost principles; (2) that the same costs that have been treated as indirect costs are not claimed as direct costs; (3) that similar types of costs, in like circumstances, have been accorded consistent accounting treatment; (4) that the information provided by the contractor/grantee, which was used as the basis for the acceptance of the rates agreed to herein and expressly relied upon by the Government in negotiating the said rates, is not subsequently found to be materially incomplete or inaccurate.

B. **ACCOUNTING CHANGES:** The rates contained in Section I of this agreement are based on the accounting system in effect at the time this agreement was negotiated. Changes to the method(s) of accounting for costs, which affects the amount of reimbursement resulting from the use of these rates, require the written approval of the authorized representative of the cognizant negotiating agency for the Government prior to implementation of any such changes. Such changes include but are not limited to changes in the charging of a particular type of cost from indirect to direct. Failure to obtain such approval may result in subsequent cost disallowances.

C. **PROVISIONAL RATES:** The provisional rates contained in this agreement are subject to unilateral amendment by the Government or bilateral amendment by the contracting parties at any time.

D. **USE BY OTHER FEDERAL AGENCIES:** The rates set forth in Section I hereof were negotiated in accordance with and under the authority set forth in 2 CFR 220. Accordingly, such rates shall be applied to the extent provided in such regulations to grants, contracts and other transactions to which 2 CFR 220 is applicable, subject to any limitations in part A of this section. Copies of this document may be provided by either party to other Federal agencies which have or intend to issue or award grants and contracts using these rates or to otherwise provide such agencies with documentary notice of this agreement and its terms and conditions.

E. **SPECIAL REMARKS:** The Government's agreement with the rates set forth in Section I is not an acceptance of the Georgia Institute of Technology Georgia Tech Research Corporation (GIT/RI)'s accounting practices or methodologies. Any reliance by the Government on cost data or methodologies submitted by GIT/RI is on a non-precedence-setting basis and does not imply Government acceptance.
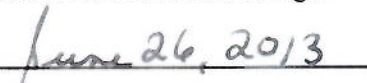
Accepted:

FOR GEORGIA INSTITUTE OF TECHNOLOGY:
GEORGIA TECH RESEARCH CORP.

_____
Jilda D. Garton
Vice President for Research,
GTRC/GTARC General Manager

_____
Date

FOR THE U.S. GOVERNMENT:

_____
for David F. Godfrey
Contracting Officer

_____
Date

For information concerning this agreement contact:
DAVID F. GODFREY, Contracting Officer
Office of Naval Research
875 North Randolph Street
Arlington, VA 22203-1995

Phone: (703) 696-2586
E-mail: david.f.godfrey@navy.mil

6. Biographies

Harley Hamilton (Ph.D., Educational Psychology, Georgia State University) Senior Research Scientist in the School of Interactive Computing at the Georgia Institute of Technology. Dr. Hamilton received the American Library Association Grolier Foundation award in 2004 for the development of the SMARTSign dictionary, an online multimedia English to American Sign Language dictionary. The ALA award is given to the person making the greatest contribution to children's literacy for the particular calendar year. He has produced video for Scholastic, Georgia Public Broadcasting, and the Georgia Department of Education. His research publications include studies in the area of psycholinguistics, sociolinguistics, education, lexicography, and dictionary use.

Scott Robertson (M.S. Computer Science, Georgia Institute of Technology) Senior Research Scientist at the Interactive Media Technology Center (IMTC) at Georgia Institute of Technology. Mr. Robertson has developed applications for distance learning, simulation, and games for training, wellness, and rehabilitation. His application development has included mobile applications, developing and directing videos, graphics and visualization, server APIs, databases, client application user experience (UX) and user interface (UI) components for various research efforts.

Don R. McCreary (Ph.D., Linguistics, University of Delaware) Professor at the University of Georgia, works primarily in lexicography and applied linguistics, as well as ESL. Dr. McCreary is the Associate Editor of the *Japanese-English Science and Engineering Dictionary* (OHM, 1988), the *English-Japanese Science and Engineering Dictionary* (OHM,1993), and the *Macmillan English Dictionary for Advanced Learners* (2002). He has authored many articles on lexicography and Japanese applied linguistics in journals such as the *International Journal of Lexicography* , *Lexicographica, Semiotica, and Language Sciences.* He is the coauthor of *Pedagogical Lexicography Today: A Critical Bibliography on Learners' Dictionaries with Special Emphasis on Language Learners and Dictionary Users* (Lexicographica Series Maior 96. Tubingen: Max Niemeyer Verlag, 1999). He also edits *"DawgSpeak,"* the dictionary of UGA student slang, which is on Georgia's English Department web page.

Dr. Daren C. Brabham (Ph.D. , Communication, University of Utah) Assistant professor in the Annenberg School for Communication & Journalism at the University of Southern California, where he conducts research and teaches in the areas of new media and public relations. Dr. Brabham was the first to publish scholarly research specifically on the topic of crowdsourcing, an online problem solving model that uses online communities to meet organizational goals. He is the author of *Crowdsourcing* (MIT Press, 2013), and his work on crowdsourcing has appeared in journals such as *Convergence*; *Information, Communication & Society*; *Planning Theory*; *First Monday*; *International Journal of Communication*; and *Journal of Applied Communication Research*. His research has been funded by the U.S. Federal Transit Administration, the IBM Center for the Business of Government, and the Social Science and Humanities Research Council of Canada, and he has built and deployed successful crowdsourcing applications for the public and private sector through research projects and consulting engagements. Brabham is the founding editor of the online, peer-reviewed journal *Case Studies in Strategic Communication.*

Plan for data management and sharing of the products of research following NSF policy on the dissemination and sharing of research results (see AAG Chapter VI.D.4). This is standard Georgia Tech policy.

1. The types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project; plans for archiving data, samples, and other research products, and for preservation of access to them.

D3 data and archiving

Data	Archived
Text of word, definition and sample sentence	In database app on GT server with the app backed up on another GT server
Video clips for words	On YouTube ("live", viewable by users) and a secure backup server at Georgia Tech (GT)
Source code for all apps	Distributed via SourceForge and D3 website, backed up on secure GT server.

2. The standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies);

Video hosted and stored in YouTube will be in the standard YouTube format to ensure cross-platform compatibility. All archived video data on Georgia Tech servers will be in mp4 format. The user apps will be created using HTML5 and Javascript/AJAX formatting . The back-end web service application programming interface (API) using a REST and JSON software architecture and the back-end database schema and data models will use an open-source database such as MongoDB or MySQL.

3. Policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements;

The user apps will be available to all to use for free.

As this project allows for crowd-produced entries and such entries are of video, the question of copyright may be encountered. We will request that contributors list the source of their video as either Crative Commons, public domain, self-made, or other. For all video, we will follow the "fair use" guidelines at http://www.teachingcopyright.org/handout/fair-use-faq which have been developed by the Electronic Frontier Foundation at Stanford University to ensure copyright compliance. We will also rely upon another safeguard to ensure our entries follow copyright laws. As the video will be YouTube based we will utilize YouTube's copyright safeguards that alert a video owner when their video may violate a copyright owner's rights. This is their Content ID solution. In December 2010, YouTube announced that more than 100 million videos had been claimed under Content ID . The video is then removed by the person who uploaded it or by YouTube. If a clip is a blatant violation of copyright law, and has been accepted in to D3

via our editorial process it will be removed from the D3 database. Our editorial process for "fair use" should be a gatekeeper to forestall any such occurrences.

4. Policies and provisions for re-use, re-distribution, and the production of derivatives; and

To ensure the ultimate educational impact for D3 entries they will all be licensed with the following Creative Commons attributes

Others are free:

- to Share — to copy, distribute and transmit the work
- to Remix — to adapt the work

Under the following conditions:

- Attribution — You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
- Noncommercial — You may not use this work for commercial purposes.

License Features

Your choices on this panel will update the other panels on this page.

Allow modifications of your work?

◉ Yes   ○ No

○ Yes, as long as others share alike

Allow commercial uses of your work?

○ Yes   ◉ No

License Jurisdiction:

International ▼

⑦ **Selected License**
Attribution-NonCommercial
3.0 Unported

(cc) (i) (S)

This is not a Free Culture License.   ✳

-

8. Letters of commitment and support

Daren C. Brabham, Ph.D.
Crowdsourcing Research & Consulting

www.darenbrabham.com
daren.brabham@gmail.com
(801) 633-4796 mobile

Dr. Harley Hamilton
Senior Research Scientist
School of Interactive Computing
Georgia Institute of Technology

July 1, 2013

Dear Dr. Hamilton,

Thank you for including me in your Dictionary 3.0 project. Please consider this letter a guarantee of my commitment as a consultant to this project, specializing in crowdsourcing.

Sincerely,

Daren C. Brabham, Ph.D.

The University of Georgia

FRANKLIN COLLEGE OF ARTS AND SCIENCES
Linguistics Program

Dr. Harley Hamilton
3.0 Dictionary Project
Georgia Tech
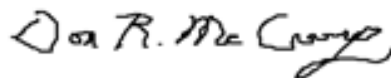Atlanta, GA

July 29, 2013

Dear Professor Hamilton,

I am very pleased to be invited to contribute to your innovative dictionary project. I look forward to consulting with you. I will provide my best advice on various details of the project.

Over the last twenty years, I have worked on both bilingual and monolingual dictionary projects. I am the Associate Editor of the *Japanese-English Science and Engineering Dictionary* (OHM, 1988), the *English-Japanese Science and Engineering Dictionary* (OHM, 1993), and I have authored many articles on lexicography in journals such as the *International Journal of Lexicography* and *Dictionaries*. I have been on the editorial advisory board for Macmillan's learners' dictionaries. I am also the coauthor of *Pedagogical Lexicography Today: A Critical Bibliography on Learners' Dictionaries with Special Emphasis on Language Learners and Dictionary Users* (Lexicographica Series Maior 96. Tubingen: Max Niemeyer Verlag, 1999).

I have been teaching Lexicography at the University of Georgia since 2001. I think the new Dictionary 3.0 project that you are developing will be innovative and pedagogically sound. I look forward to working on it.

Sincerely,

Don R. McCreary,
Professor of Linguistics and English
Email: mccreary@uga.edu

August 1, 2013

Scott L. Robertson
Senior Research Scientist
Interactive Media Technology Center (IMTC)
Georgia Institute of Technology
85 5th Street NW, Rm 314
Atlanta, Georgia 30308

Dear Dr. Hamilton,

I am writing to enthusiastically endorse your proposal for NEH support for the Dictionary 3.0 Project. I also wish to confirm my own, as well IMTC's, interest in collaborating on the research proposed. As you know, our previous collaboration on research projects such as the mobile device and desktop American Sign Language (ASL) video dictionaries will provide a solid foundation to build upon for this new research effort.

One of IMTC's strengths is developing robust software systems, which employ industry best practices for user, interface (UI) and user experience (UX) design, particularly for mobile device applications. Our nationally recognized work in the areas of training, education and e-learning will also inform this new collaboration. I and IMTC will be pleased to be actively involved in this research, providing insight and advice on system architecture and mobile application user experience and usability issues as well as designing, developing and deploying the front and back end software for the Dictionary 3.0 system.

I look forward to this cooperation with you and your colleagues.

Sincerely,

Scott L. Robertson
Senior Research Scientist
Interactive Media and Technology Center
School of Interactive Computing
Georgia Institute of Technology

Banks County School System

CHRISTOPHER B. ERWIN
Superintendent

DONNA F. REED
Assistant Superintendent

May 14, 2013

To Whom It May Concern,

I am writing this letter in support of the Dictionary 3.0 grant proposal. This dictionary will be a valuable tool for students and teachers. I will gladly allow our students to participate in field studies to ensure the development of this tool. I look forward to the day when the meaning of English print is quickly accessible to readers and hope you will fund the Dictionary 3.0 project as a step in this direction.

Sincerely,

Christopher B. Erwin

CHEROKEE COUNTY SCHOOL DISTRICT
OFFICE OF THE SUPERINTENDENT

May 28, 2013

TO: Dr. Harley Hamilton
              Senior Research Scientist, School of Interactive Computing
              Georgia Institute of Technology

FROM: Dr. Frank R. Petruzielo
              Superintendent of Schools

SUBJECT: GEORGIA INSTITUTE OF TECHNOLOGY PARTNERSHIP

I am writing to express the Cherokee County School District's support of your Dictionary 3.0 grant proposal and this opportunity for partnership with the Georgia Institute of Technology.

The School District will allow Clark Creek Elementary School STEM Academy students to participate in field studies to ensure the development of this tool.

We look forward to seeing this dictionary become reality, as it would be a valuable tool for all readers – especially students and teachers; and we hope you will gain the grant funding you seek to make the Dictionary 3.0 project possible.

_____FRP

cc: Superintendent's Cabinet
        Dr. Jennifer Scrivner, Principal, Clark Creek ES STEM Academy

Lamar County Schools

100 Victory Lane
Barnesville, Georgia 30204-1544

770.358.5891
Superintendent

Fax: 770.358.5858

To: Whom it may concern

From: William F. Truby, PhD
      Superintendent of Schools, Lamar County, GA

Date: May 14, 2013

RE: Dictionary 3.0 Grant

I have been made aware of the opportunity to participate in a hands-on

Dictionary 3.0 Grant for our students in the Lamar County School System.

Therefore, I am writing this letter in support of the Dictionary 3.0 grant

proposal. This dictionary will be a valuable tool for students and teachers.

I will gladly allow our students to participate in field studies to ensure the

development of this tool. I look forward to the day when the meaning of

English print is quickly accessible to readers and hope you will fund the

Dictionary 3.0 project as a step in this direction.

9. **Appendices**

Appendix A- Web-based submission form (return to document)

Contributor name (optional) _____ Email address _____ Word _____ Definition _____ Sample sentence _____ Video url _____

Appendix B- Acceptance criteria ([return to document](#))

To be accepted the entry must have

- o A clear, accurate definition
- o a simple sample sentence which is clearly illustrated by video
- o video or images appropriately licensed via Creative Commons, in the public domain, or fit the terms of "fair use". We will follow the "fair use" guidelines at http://www.teachingcopyright.org/handout/fair-use-faq which have been developed by the Electronic Frontier Foundation at Stanford University.
- o a rating by D3 editors as G or PG following the rating system described by Motion Picture Association of America

G — General Audiences. All Ages Admitted. A G-rated motion picture contains nothing in theme, language, nudity, sex, violence or other matters that, in the view of the Rating Board, would offend parents whose younger children view the motion picture. The G rating is not a "certificate of approval," nor does it signify a "children's" motion picture. Some snippets of language may go beyond polite conversation but they are common everyday expressions. No stronger words are present in G-rated motion pictures. Depictions of violence are minimal. No nudity, sex scenes or drug use are present in the motion picture.

PG — Parental Guidance Suggested. Some Material May Not Be Suitable For Children. A PG-rated motion picture should be investigated by parents before they let their younger children attend. The PG rating indicates, in the view of the Rating Board, that parents may consider some material unsuitable for their children, and parents should make that decision. The more mature themes in some PG-rated motion pictures may call for parental guidance. There may be some profanity and some depictions of violence or brief nudity. But these elements are not deemed so intense as to require that parents be strongly cautioned beyond the suggestion of parental guidance. There is no drug use content in a PG-rated motion picture. (D3 will scrutinize these videos but considers them necessary for words such as "murder" . Graphic depictions will not be accepted.)

References

Afflerbach, P., Beers, J., Boyd, C.D., Blachowicz, C., & Diffily, D. (2000). Reading : Surprise Me / Edition 1 . Reading : Surprise Me / Edition 1. Boston: Addison-Wesley Educational Publishers, Inc.

Alloway, T.P., Gathercole, S.E., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development*, 80(2), 606–621**.**

Beech, J.R. (2004). Using a dictionary: its influence on children's reading, spelling, and phonology**.** *Reading Psychology*, 25,19–36.

Borst, I. (2010). Understanding crowdsourcing; Effects of motivation and rewards on participation and performance in online voluntary activities. PhD thesis Erasmus Research Institute of Management, Erasmus University.

Brabham, D. C. (2012). Crowdsourcing: A model for leveraging online communities. In A. Delwiche & J. J. Henderson (Eds.), *The Participatory Cultures Handbook. (pp.120-129). UK:*Routledge.

Denh, M, (2008) *Working memory and academic learning: assessment and intervention*. Hoboken, NJ: John Wiley & Sons.

De Schryver, G-M. 2003. Lexicographers' Dreams in the Electronic-Dictionary Age. *International Journal of Lexicography* 16(2): 143-199.

Fillmore, C.J. (2003). Double-Decker Definitions: The role of frames in meaning explanations Sign Language Studies, 3(3), Spring 2003, 263-295.

Hamilton, H. (2012). The efficacy of dictionary use while reading for learning new words. *American Annals of the Deaf,* 157(4), 358-372.

Johnson, S. (1755). *A Dictionary of the English Language*. London: J. & P. Knapton.

Lew, R. (2010). New ways of indicating meaning in electronic dictionaries: Hope or hype? In Y. Zhang, (Ed.), *Learner's Lexicography and Second Language Teaching*. (pp. 387-404.Shanghai: Shanghai Foreign Language Education Press.

Weld, D.S., Adar, E., Chilton, L., Hoffmann, R., Horowitz, E., Koch, M., Landay, J., Lin, C.H. & Mausam. (2012). Personalized Online Education -A Crowdsourcing Challenge. *Human Computation AAAI Technical Report*, WS-12-08.

Winchester, S. (1998). *The Professor and the Madman: A Tale of Murder, Insanity, and the Making of the Oxford English Dictionary*, New York: Harper Perennial.

Study 1. Usability of D3 interface.

Research question- What interface features allow students to easily use D3 to access word meaning?

Participants- Ten 3rd grade children, ten 4th grade children, ten 5th grade children. This is the age where academic vocabulary begins to outpace a student's oral vocabulary and the need for a vocabulary tool increases.

Procedure- Students will be shown given a list of ten words and asked to find the meaning of each using D3. They will be shown how to use D3 to find word meanings. They will be retested one week after the original session to measure retention of D3 usage procedures. We will measure usability by observation based on the Info-Tech usability instrument designed by the Info-Tech research group http://www.infotech.com/.

Analysis- Determine usability based on score from observational data. Scores will be tallied for each word the students look up and averaged for factors 1-4. Factors 5 -7 will be one global score for a session. If the average score for a goal is below 4.5/5 that feature of D3 will be revised and retested. This study will be iterated as needed to obtain a D3 interface which is easily usable by children in grades 3-5.

Info-Tech Usability scale

Factor	Ranking (1= low, 5 = high)	Measurable Requirements
1.Automation- A high ranking means that much of the task is routine and the user should not be bothered by it.		Students will quickly and easily look at the word on the list and begin to enter it in the D3 text input box. This replicates the student looking at a weekly vocabulary list and using a dictionary to find definitions, a common school task
2.Efficiency (Fast)- A high ranking means that the user can quickly and efficiently use D3 to access word meanings		Students will easily enter target word in text input box using the keyboard.
3.Learnability- A high ranking means that learning is not a major issue. A low ranking means learning the system is an important issue		After a brief introduction on how to use D3 we will observe whether the students easily follow the procedure of 1. Look at word on list 2. Enter word in D3 3. Explore the dictionary entry During the third step, it will be noted how many times the child plays a video and listens to the definition and

		sample sentence via use of the "speak" buttons, and amount of time they spend on each entry..
4.Reliability (Protection Against User Errors)- A high ranking means that errors or distractions are minimal and can be tolerated in order to complete task. A low ranking means that there are errors and distractions possible and that users should be protected from them either by preventing them or by making recovery effective.		Is text entry an efficient means of initiating a query?
5.Satisfying Subjectively- A high ranking means that, assuming all the other priority criteria are acceptable, using D3 is a pleasant experience for the user. A low ranking means that users need not like the system. The questions will be scored at one point a piece with fractions as possible scores. For example, a response to the first question of "yes" = 1 point, "kinda" or it's OK=1/2 point, and "no"= 0.		We will ask the students how they liked using the dictionary? What was their favorite word? Do they want to use it again? Was there any hard part to using the dictionary? Do they want to use it again
6.Understandability- A high ranking means the users grasp the concepts of how D3 works and how they can use it. A low ranking means that the user does not see the use of D3 as a learning tool. Each question will be valued at 2.5 points.		The students will be asked when they would use D3 if they had it in school? The students will be asked when they would use D3 if they had it at home?
7.Memorability- A high ranking means that users will use the system, be away from it for extended periods, then return to use it (i.e. intermittent use). It will be important for them to get back up to speed quickly.		We will bring the students back one week after the initial session and replicate the experiment using a different set of words.

Study 2.
Research question- Are the crowdsourcing and submission mechanism instructional videos and process easy for users to understand and complete?
Participants- Ten individuals age 18 or over.
Procedure- Participants will view the D3 instructional videos demonstrating how to create and submit D3 entries. They will then be asked to create and submit five entries using the YouTube editor and Creative Commons media for the words, *cat*, *run*, *fall, sad*, and *spooky*. They will then complete the following Likert scale regarding the video creation and entry submission task in order to determine the overall ease of these tasks and determine areas which need to be modified to make the task manageable. Comments will also be collected
Analysis- Mean scores will be calculated for each statement as well as an overall task mean. If scores are below 4, iteration will be done to modify the particular task. Comments will also guide any modifications.

Making videos- Circle your response

1. Understanding the video creation tutorial was

very hard	hard	manageable	easy	very easy
1	2	3	4	5

Comments _____

2. Creating video with the YouTube editor was

very hard	hard	manageable	easy	very easy
1	2	3	4	5

Comments _____

Submitting videos- Circle your response

1. Understanding the video submission tutorial was

very hard	hard	manageable	easy	very easy
1	2	3	4	5

Comments _____

2. Submitting videos was

very hard	hard	manageable	easy	very easy
1	2	3	4	5

Comments _____

1. **Table of contents**

2. **List of participants**
Harley Hamilton- Georgia Institute of Technology
Scott Robertson- Georgia Institute of Technology
Don McCreary- University of Georgia
Daren Brabham- University of Southern California

3. **Abstract**

Dictionary 3.0 (D3) will be a free multimedia English language dictionary that provides a powerful reference tool for students, adult second language learners, and life-long learners, essentially anyone. As a humanities reference tool it will be an invaluable living repository of the English lexicon as it will be continually grow and be updated via crowdsourcing. D3 will revolutionize the presentation format of dictionaries which has changed little since Samuel Johnson's *A Dictionary of the English Language* (1755). Even today, text dominates as the presentation medium, yet technology allows much more. D3's multimedia presentation will include not only text to represent meaning but text-to-speech capabilities, audio, and video to illustrate word meaning in a rich context, just as word meaning is embodied in everyday language usage. <u>Meaning</u> is what dictionary users most desire (Lew 2010). D3 will deliver word meaning that is accessible on any device, computer, tablet, phone…

<center>"Any word, anywhere, anytime"</center>

Statement of Innovation
New approaches to dictionary-making practice are needed as the field of lexicography seeks to invigorate itself (Alonso, Millon, & Williams, 2011). Dictionary design today should be dominated by "innovation and bold intelligence" not tradition (Gouws (2011). A dictionary should be a monofunctional tool focused solely on providing word meaning (Leroyer, 2011; Berghenholtz & Bergenholtz, 2011). D3 will be such an innovative and boldly intelligent tool.

Statement of Humanities Significance
The humanities include dictionary creation, language study, and history of the arts. D3 will provide context rich examples of the English lexicon via video clips that may serve linguistic study of usage. As clips will be culled from video produced over the last 100 years it will also provide scholars with a historical database of language usage and preserve examples of the visual and performing arts over that time. Crowdsourcing will allow for everyman to contribute to this humanities reference.

4. **Narrative**

o **Enhancing the humanities through innovation**

We are seeking a Phase I, Level II grant to plan and develop a prototype of D3. This work will lead to a functional set of apps and tutorials that will be a complete infrastructure for the fulfillment of D3 in the next phase grant which will include creation of the D3 lexicon by grant staff and the public via crowdsourcing. The creation of D3 will revolutionize the field of dictionary making while providing an educational tool to assist students with language learning and reading. It will also provide a participatory humanities activity (crowdsourcing) that allows anyone to contribute to the creation of a humanities reference work containing the lexicon of English. The creation of D3 entries by the general public can be accomplished by open-source/freely accessible tools. People will be able to create contributions in several ways,

1. Using a digital camera, camcorder, or phone to capture video and post it to YouTube. These devices are common and owned by a vast majority of individuals.
Their video can be edited in several free editing programs such as the YouTube editor, Windows Movie Maker, and iMovie as well as other free options
2. Using Creative Commons media which are integrated with the free YouTube editing app to create new videos. There are over 4 million videos and hundreds of millions of still images available with Creative Commons licensing that allows for re-use.
3. Free tools such as Muvizu, FreshBrain, and WeGame are available for making animated videos.

Workflow
Contributor tasks
- Establish own YouTube channel.
- Create entry and upload to YouTube channel
- Submit entry to D3 via web form (see Appendix A)

D3 editorial tasks- performed initially by project director for user studies until volunteer editorial network is developed as part of later grant development and funding.
- Review submitted entry to accept, reject, or comment for revision and resubmission
  - o If accepted, (see Appendix B for acceptance criteria)
    - Accepted video is uploaded to D3's YouTube channel
    - The word, definition, sample sentence, and video url are entered in the D3 database for access via D3 user app.

- o If rejected, contributor is informed of the D3 decision and the reason for rejection
- o If revisions are suggested the entry will then be reconsidered upon resubmission

Crowdsourcing is an economical and common method for dictionary development. The sheer magnitude of a project capable of producing a dictionary covering the words of a language precludes the possibility of any single publisher creating such a tool due to time and cost-restraints (Fillmore, 2003). This is even more pertinent when all entries contain audio and video, not simply text. Scaling such a tool requires a crowd (Weld et al. 2012) and has been successful in the past. In 1859, the British Philological Society launched an appeal to the British and American public to create a new dictionary "worthy of the English Language and of the present state of philological science." Over 6 million submissions were made over a period of 70 years (Winchester, 1998). This helped produce the Oxford English Dictionary. Today, the Oxford English dictionary is still employing this tactic. Other dictionary makers also employ crowdsourcing

http://www.macmillandictionary.com/us/open-dictionary/

http://nws.merriam-webster.com/opendictionary/submit.php

http://www.ahdictionary.com/word/submitNew.html

D3 follows in their footsteps to create a dictionary worthy of the English language and the present state of technological as well as philological science. To facilitate crowd-sourcing, we will rely on both intrinsic and extrinsic motivations. These include fun, learning, recognition, self-satisfaction from contributing to a large project of common interest, altruism, and creative expression (Borst, 2010; Brabham, 2012). D3 is a hybrid Wiktionary and YouTube entity, both of which are populated by contributions by users enticed by the motivations just mentioned. Contributing to YouTube is a rather easy task as evidenced by the sheer volume of video contributed. One hundred hours of video are uploaded to YouTube each minute.

o **Environmental scan**

Dictionaries of today
  Dictionary 1.0 (D1)
  Before the technological revolution and birth of the internet, dictionaries were massive text-based paper tomes. These still exist today just as the coelacanth remains extant (see The American Heritage Student Dictionary, 2012). "Illustrated" dictionaries include a few images but remain text-dominant (see Merriam-Webster Elementary Dictionary, 2010). "Picture" dictionaries for younger students are available but the content is generally limited to easily imaged content such as nouns, action verbs, and some adjectives (see Scholastic First Picture Dictionary, 2009). Words such as *spooky* (found in a 1st grade reader, Afflerbach, et al., 2000) would be difficult if not impossible to clearly illustrate with a static image and are not included. However, the word *blue tit* (a bird) is. Each of the aforementioned dictionaries is rated as 5 out of 5 stars on Amazon.com.
  Finding words in paper dictionaries is also problematic and time consuming as a fairly high level of alphabetizing skills is needed. Studying children age 7-11, Beech (2004) found that

for younger readers (age 7-9.4) the average look up time for a printed word was 48 seconds with an accuracy rate of 68%. For older children (age 9.5-11) the average look up time was 28 seconds with an average accuracy rate of 90%. Clearly, for many children, finding words in a paper dictionary is time consuming and an inefficient, inaccurate means of accessing word meaning. The search time plus encountering unknown words in the definitions themselves take attention and memory resources away from the reading passage for a rather long period of time and may overload working memory and negatively impact reading comprehension (Denh, 2008; Alloway, Gathercole, Kirkwood, & Elliott, 2009).

Dictionary 2.0 (D2)

D2 or electronic dictionaries may be web-based, on CD/DVDs, or stand-alone devices. The limitations of the early internet (Web 1.0) restricted D2 to mainly text-based entries. CD/DVDs allowed for multimedia entries but were still limited due to the large size of video and image files and the relatively small storage size of the disks themselves. CD/DVDs are also limited in ease of accessibility due to the need for a computer to utilize them. These media are also not cost effective as they range in price from $10 to over $200 . Stand-alone devices such as those made by Franklin range in price from $50 to $400. They are predominantly text-based due to technology limitations and cost.

As technology developed and the internet evolved, D2 began to appear online. As Web1.0 became Web 2.0 with full multimedia capabilities dictionary development, however, has remained stagnant. An online dictionary may be just a copy of the paper dictionary (see www.webster-dictionary.org/). The top two online adult dictionaries as determined by a Google search (see www.merriam-webster.com/, and www.dictionary.com) do include a minor multimedia component that allows the user to hear only the target word spoken. The user interface of each of these is extremely cluttered with ads and other features such as links to quizzes, word of the day, a thesaurus, Spanish…. These two sites also include sample sentences. However, the sample sentences are inadequate unless you already know the meaning of the target word. For example at www.dictionary.com a sample sentence for *spooky* is "They make trees look spooky, and they prefer not to be disturbed." The enigmatic nature of this sentence is overwhelming, perhaps even spooky.

Visual dictionaries have appeared, yet are limited to static images, primarily of nouns. The breadth of the largest online visual dictionary (http://visual.merriam-webster.com/ is 20,000 words and 6,000 images. *Spooky* cannot be found. This may cause users frustration due to lack of a substantial lexicon. Visual webs such as http://www.visualthesaurus.com/ are also available yet remain exclusively text-based.

Today, most web-based dictionaries remain text-based or include only minor multimedia components. While access time and accuracy of search may be improved by allowing the user to type the target word in a text box, current on-line dictionaries do not take full advantage of the potential offered by Web 2.0 technology (deSchryver, 2003). Such a statement was true a decade ago and remains accurate to this day.

Dictionary 2.5 (D2.5)

Some attempts are being made to utilize the multimedia power of Web 2.0. Wordia primarily utilizes video of people reading the definition or explaining the word alongside its text definition. Meaning is primarily rendered by the spoken word. It is crowdsourced and requests users to upload entries when video related to a text definition does not exist. Its editorial policy

appears to allow for esoteric tributes rather than clear representations of meaning. See a definition for _dog_ contributed by nickgrim (lower left hand corner of web page).

Vidtionary contains video and is limited to easily represented nouns and verbs (although the word _cat_ could not be found). Their video definitions can be confusing, search for _color_ to see an example. Once again _spooky_ is unavailable in each, even though the media lends itself to a vivid depiction of the meaning of _spooky_.

Dictionary 3.0 (D3)

D3 will take advantage of the multimedia capabilities of Web 2.0, the popularity of inexpensive devices such as phones, tablets, and laptops, and the fully connected nature of today's technological infrastructure to become an inexpensive, media-rich, ubiquitous reference work on the English lexicon that brings word meaning to those who need it "any word, anywhere, anytime". Word meaning will be illustrated via video rich in physical and social context clearly depict meaning, just as in everyday life. See _spooky_ . The meaning of _spooky_ in this example is illustrated by both the video and its accompanying spooky audio as well as the definition and sample sentence which are directly related to the video.

o **History and duration of the project**

D3 began in January, 2013. It is the evolution of the project director's work on the English to American Sign Language bilingual dictionary, SMARTSign. Work on the SMARTSign dictionary was awarded the American Library Association's Grolier Foundation award for making the greatest contribution to children's literacy in 2004. A review of the literature for a study of the efficiency of dictionary use by deaf signing students (Hamilton, 2012) made the project director aware of the impact of vocabulary on reading for all readers, hearing and deaf. This review also made it clear that while dictionaries are powerful learning tools they are sorely underutilized due to their current format. That information along with knowledge of what current technology can provide melded into the idea of D3. When demos have been done for a wide range of consumers, the general response has been "That is so cool!" (even by people my age, 60). Enlisting collaborators from academia has been easy and elicited the same response.

D3 currently has no financial support. As a research scientist at Georgia Tech, the project director is seeking funding for this project from NEH and other governmental sources. We will continue to seek funding to make D3 a living, growing, social media platform that provides an avenue for participants to collectively create a complete reference work of the English lexicon as well as a powerful learning tool for students.

o **Work plan**

This work will include development of

- D3 back-end database. Design back-end database schema and data models using an open-source database such as MongoDB or MySQL. The database schema will include fields for storing dictionary terms, their associated URL links, text definition and sample sentences, user interaction logging including time-stamped entries per user, dictionary terms searched for, and subsequent hits and misses. Person responsible- Scott Robertson
- Server API design: Design and implement a back-end web service application programming interface (API) using a REST and JSON software architecture. This API will allow client applications (apps) to create, read, update and delete dictionary database

records as well as create and read user interaction log records. This API will support an administrative web app, used to administer the dictionary database (i.e. add new entries), client web and mobile apps and desktop browser widgets or extensions. Person responsible- Scott Robertson

- Client mobile app design: Design and implement a cross-platform, HTML5 and Javascript/AJAX based mobile app and user interface to allow users to search for dictionary terms and view the targeted entry. This app will be developed using a development framework such as Sencha Touch in order to run on multiple platforms: iOS, Android, Blackberry, Windows Mobile. Such HTML5 apps can be packaged and delivered as native device apps or can be run in a device's mobile web browser. Person responsible- Scott Robertson

- Desktop browser widget and/or extension: Design and develop a web browser extension and user interface which will implement the following functionality: Allow a user to highlight any word on a web page she is browsing and perform a D3 dictionary search, presenting the search result video(s) in a pop-up browser window. This extension will initially be developed for the Google Chrome browser using the Google Chrome extensions API. Google Chrome is a cross-platform browser and Chrome extensions can be used on Windows and Apple OSX platforms. Person responsible- Scott Robertson

- Crowdsourcing app- A web-based form will be developed using HTML and javascript that will allow people to submit entries to D3 for acceptance or editorial suggestions. This will include word, definition, sample sentence, video url, and contact information. Contributors will store their videos on their own YouTube channel. If an entry is accepted it will then become part of D3 and stored on the D3 YouTube channel to ensure availability. Person responsible- Scott Robertson

- D3 lexicon- Identify scope of D3 lexicon and provide a web-based resources for suggested entries for crowdsourcing participants. Persons responsible- Don McCreary, Harley Hamilton

- Sample D3 entries- Five hundred D3 entries will be created. Person responsible- Harley Hamilton

- Video tutorials- YouTube-based tutorials showing various ways to create and submit D3 entries using several different free tools such as Windows Movie Maker, iMovie, Muvizu, WeGame, YouTube video editor, Creative Commons content, and video camera. Person responsible- Harley Hamilton

- Crowdsourcing mechanism- Develop means to inform the public of D3 and encourage participation in its creation as well as the means for submitting entries. Person responsible- Daren Brabham, Harley Hamilton

- Update D3 webpage- As D3 develops, its webpage will be used as a major source of new information for social and mainstream media outlets. Person responsible- Harley Hamilton

- User studies- Studies to iteratively refine the interface, crowdsourcing mechanism, and tutorials. These will inform the development of D3. See attached studies. IRB clearance will be obtained via normal Georgia Tech channels. This entails completion of a detailed IRB request made online to the Office of Research Integrity Assurance http://researchintegrity.gatech.edu/. Letters of support from local school districts to participate in the studies are in appendix C. Person responsible- Harley Hamilton

Sustainability

- Technology- All apps will reside and be backed up on secure servers at Georgia Tech. The videos will be hosted by YouTube via the D3 YouTube channel. Backup copies of all video will be stored on secure Georgia Tech servers.
- D3 editing and updating- If the initial grant is funded, a subsequent full grant will be applied for. During the second grant period, funds will be requested for two research assistants whose responsibilities will include the editorial task for accepting entries and including these in the D3 lexicon. If an entry passes both assistants, the project director will have the final say on an entry's acceptance. During this second grant, a crowdsourced editorial board will be assembled following the lead of Wikipedia and Wiktionary to ensure continued D3 growth after funding is no longer available.

Staff

Harley Hamilton, (Project director)- Dr. Hamilton will oversee the project as well as create sample entries and tutorials, and conduct user studies to iteratively refine D3. He will work on D3 1.5 months over the one year grant period.

Scott Robertson, (Lead developer)- Mr. Robertson will develop all apps for D3 as described above. He will devote two months to D3 over the course of the grant.

Don McCreary, (Dictionary consultant)- With Dr. Hamilton, Dr. McCreary will formulate the initial lexicon to be crowdsourced in order to populate D3.

Daren Brabham, (Crowdsourcing consultant)- Dr. Brabham will design the plan and mechanisms for implementing D3 crowdsourcing.

o **Final product and dissemination**

Social media will be used to generate "buzz" so that future crowdsourcing can hit the ground running. D3 already has Twitter, YouTube, and Facebook accounts established and these will be utilized to promote it. Academic articles describing the development of D3 will be submitted to the International Journal of Lexicography and Lexicographica. Press releases will be sent to all major news outlets via the Georgia Tech media relations department. The D3 webpage will be updated and used to refer people to via tweets and press releases.

4. **Project budget – Attached**

   Budget narrative (optional)

6. **Biographies**

Harley Hamilton (Ph.D., Educational Psychology, Georgia State University) Senior Research Scientist in the School of Interactive Computing at the Georgia Institute of Technology. Dr. Hamilton received the American Library Association Grolier Foundation award in 2004 for the development of the SMARTSign dictionary, an online multimedia English to American Sign Language dictionary. The ALA award is given to the person making the greatest contribution to

children's literacy for the particular calendar year. He has produced video for Scholastic, Georgia Tech, Georgia Public Broadcasting, and the Georgia Department of Education. His research publications include studies in the area of psycholinguistics, sociolinguistics, education, lexicography, and dictionary use. He researched and published seven American Sign Language dictionaries.

Scott Robertson (M.S. Computer Science, Georgia Institute of Technology) Senior Research Scientist at the Interactive Media Technology Center (IMTC) at Georgia Institute of Technology. Mr. Robertson has developed applications for distance learning, simulation, and games for training, wellness, and rehabilitation. His application development has included mobile applications, developing and directing videos, graphics and visualization, server APIs, databases, client application user experience (UX) and user interface (UI) components for various research efforts.

Don R. McCreary (Ph.D., Linguistics, University of Delaware) Professor at the University of Georgia, works primarily in lexicography and applied linguistics, as well as ESL. Dr. McCreary is the Associate Editor of the *Japanese-English Science and Engineering Dictionary* (OHM, 1988), the *English-Japanese Science and Engineering Dictionary* (OHM,1993), and the *Macmillan English Dictionary for Advanced Learners* (2002). He has authored many articles on lexicography and Japanese applied linguistics in journals such as the *International Journal of Lexicography* , *Lexicographica, Semiotica, and Language Sciences.* He is the coauthor of *Pedagogical Lexicography Today: A Critical Bibliography on Learners' Dictionaries with Special Emphasis on Language Learners and Dictionary Users* (Lexicographica Series Maior 96. Tubingen: Max Niemeyer Verlag, 1999). He also edits *"DawgSpeak,"* the dictionary of UGA student slang, which is on Georgia's English Department web page.

Dr. Daren C. Brabham (Ph.D., Communication, University of Utah) Assistant professor in the Annenberg School for Communication & Journalism at the University of Southern California, where he conducts research and teaches in the areas of new media and public relations. Dr. Brabham was the first to publish scholarly research specifically on the topic of crowdsourcing, an online problem solving model that uses online communities to meet organizational goals. He is the author of *Crowdsourcing* (MIT Press, 2013), and his work on crowdsourcing has appeared in journals such as *Convergence*; *Information, Communication & Society*; *Planning Theory*; *First Monday*; *International Journal of Communication*; and *Journal of Applied Communication Research*. His research has been funded by the U.S. Federal Transit Administration, the IBM Center for the Business of Government, and the Social Science and Humanities Research Council of Canada, and he has built and deployed successful crowdsourcing applications for the public and private sector through research projects and consulting engagements. Brabham is the founding editor of the online, peer-reviewed journal *Case Studies in Strategic Communication.*

7. **Data management plan**
Following Georgia Tech policy we have developed a data management plan following the National Science Foundation guidelines for such a plan. See attached.

8. Letters of commitment and support

Daren C. Brabham, Ph.D.
Crowdsourcing Research & Consulting

www.darenbrabham.com
daren.brabham@gmail.com
(801) 633-4796 mobile

Dr. Harley Hamilton
Senior Research Scientist
School of Interactive Computing
Georgia Institute of Technology

July 1, 2013

Dear Dr. Hamilton,

Thank you for including me in your Dictionary 3.0 project. Please consider this letter a guarantee of my commitment as a consultant to this project, specializing in crowdsourcing.

Sincerely,

Daren C. Brabham, Ph.D.

The University of Georgia
FRANKLIN COLLEGE OF ARTS AND SCIENCES
Linguistics Program

Dr. Harley Hamilton
3.0 Dictionary Project
Georgia Tech
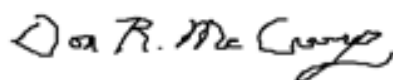Atlanta, GA

July 29, 2013

Dear Professor Hamilton,

I am very pleased to be invited to contribute to your innovative dictionary project. I look forward to consulting with you. I will provide my best advice on various details of the project.

Over the last twenty years, I have worked on both bilingual and monolingual dictionary projects. I am the Associate Editor of the *Japanese-English Science and Engineering Dictionary* (OHM, 1988), the *English-Japanese Science and Engineering Dictionary* (OHM, 1993), and I have authored many articles on lexicography in journals such as the *International Journal of Lexicography* and *Dictionaries*. I have been on the editorial advisory board for Macmillan's learners' dictionaries. I am also the coauthor of *Pedagogical Lexicography Today: A Critical Bibliography on Learners' Dictionaries with Special Emphasis on Language Learners and Dictionary Users* (Lexicographica Series Maior 96. Tubingen: Max Niemeyer Verlag, 1999).

I have been teaching Lexicography at the University of Georgia since 2001. I think the new Dictionary 3.0 project that you are developing will be innovative and pedagogically sound. I look forward to working on it.

Sincerely,

Don R. McCreary,
Professor of Linguistics and English
Email: mccreary@uga.edu

Georgia Tech | College of Computing
School of Interactive Computing

August 1, 2013

Scott L. Robertson
Senior Research Scientist
Interactive Media Technology Center (IMTC)
Georgia Institute of Technology
85 5th Street NW, Rm 314
Atlanta, Georgia 30308

Dear Dr. Hamilton,

I am writing to enthusiastically endorse your proposal for NEH support for the Dictionary 3.0 Project. I also wish to confirm my own, as well IMTC's, interest in collaborating on the research proposed. As you know, our previous collaboration on research projects such as the mobile device and desktop American Sign Language (ASL) video dictionaries will provide a solid foundation to build upon for this new research effort.

One of IMTC's strengths is developing robust software systems, which employ industry best practices for user, interface (UI) and user experience (UX) design, particularly for mobile device applications. Our nationally recognized work in the areas of training, education and e-learning will also inform this new collaboration. I and IMTC will be pleased to be actively involved in this research, providing insight and advice on system architecture and mobile application user experience and usability issues as well as designing, developing and deploying the front and back end software for the Dictionary 3.0 system.

I look forward to this cooperation with you and your colleagues.

Sincerely,

Scott L. Robertson
Senior Research Scientist
Interactive Media and Technology Center
School of Interactive Computing
Georgia Institute of Technology

P.O. Box 248
Homer, GA 30547
Phone: 706-677-2224
Fax: 706-677-2223

Banks County School System

CHRISTOPHER B. ERWIN
Superintendent

DONNA F. REED
Assistant Superintendent

May 14, 2013

To Whom It May Concern,

I am writing this letter in support of the Dictionary 3.0 grant proposal. This dictionary will be a valuable tool for students and teachers. I will gladly allow our students to participate in field studies to ensure the development of this tool. I look forward to the day when the meaning of English print is quickly accessible to readers and hope you will fund the Dictionary 3.0 project as a step in this direction.

Sincerely,

Christopher B. Erwin

<div style="border: 2px solid black; padding: 10px;">

**CHEROKEE COUNTY SCHOOL DISTRICT
OFFICE OF THE SUPERINTENDENT**

</div>

May 28, 2013

TO: Dr. Harley Hamilton
          Senior Research Scientist, School of Interactive Computing
          Georgia Institute of Technology

FROM: Dr. Frank R. Petruzielo
          Superintendent of Schools

SUBJECT: **GEORGIA INSTITUTE OF TECHNOLOGY PARTNERSHIP**

I am writing to express the Cherokee County School District's support of your Dictionary 3.0 grant proposal and this opportunity for partnership with the Georgia Institute of Technology.

The School District will allow Clark Creek Elementary School STEM Academy students to participate in field studies to ensure the development of this tool.

We look forward to seeing this dictionary become reality, as it would be a valuable tool for all readers – especially students and teachers; and we hope you will gain the grant funding you seek to make the Dictionary 3.0 project possible.

_____FRP

cc: Superintendent's Cabinet
          Dr. Jennifer Scrivner, Principal, Clark Creek ES STEM Academy

Lamar County Schools

100 Victory Lane
Barnesville, Georgia 30204-1544

770.358.5891
Superintendent

Fax: 770.358.5858

To: Whom it may concern

From: William F. Truby, PhD
Superintendent of Schools, Lamar County, GA

Date: May 14, 2013

RE: Dictionary 3.0 Grant

I have been made aware of the opportunity to participate in a hands-on Dictionary 3.0 Grant for our students in the Lamar County School System. Therefore, I am writing this letter in support of the Dictionary 3.0 grant proposal. This dictionary will be a valuable tool for students and teachers. I will gladly allow our students to participate in field studies to ensure the development of this tool. I look forward to the day when the meaning of English print is quickly accessible to readers and hope you will fund the Dictionary 3.0 project as a step in this direction.

The mission of Lamar County Schools is to work with and encourage our families and community to provide and participate in world class educational opportunities that inspire all learners to reach their potential.

9. **Appendices**

Appendix A- Web-based submission form ([return to document](#))

Contributor name (optional) _____ Email address _____ Word _____ Definition _____ Sample sentence _____ Video url _____

Appendix B- Acceptance criteria ([return to document](#))

To be accepted the entry must have

- A clear, accurate definition
- a simple sample sentence which is clearly illustrated by video
- video or images appropriately licensed via Creative Commons, in the public domain, or fit the terms of "fair use". We will follow the "fair use" guidelines at http://www.teachingcopyright.org/handout/fair-use-faq which have been developed by the Electronic Frontier Foundation at Stanford University.
- a rating by D3 editors as G or PG following the rating system described by Motion Picture Association of America

G — General Audiences. All Ages Admitted. A G-rated motion picture contains nothing in theme, language, nudity, sex, violence or other matters that, in the view of the Rating Board, would offend parents whose younger children view the motion picture. The G rating is not a "certificate of approval," nor does it signify a "children's" motion picture. Some snippets of language may go beyond polite conversation but they are common everyday expressions. No stronger words are present in G-rated motion pictures. Depictions of violence are minimal. No nudity, sex scenes or drug use are present in the motion picture.

PG — Parental Guidance Suggested. Some Material May Not Be Suitable For Children. A PG-rated motion picture should be investigated by parents before they let their younger children attend. The PG rating indicates, in the view of the Rating Board, that parents may consider some material unsuitable for their children, and parents should make that decision. The more mature themes in some PG-rated motion pictures may call for parental guidance. There may be some profanity and some depictions of violence or brief nudity. But these elements are not deemed so intense as to require that parents be strongly cautioned beyond the suggestion of parental guidance. There is no drug use content in a PG-rated motion picture. (D3 will scrutinize these videos but considers them necessary for words such as "murder" . Graphic depictions will not be accepted.)

References

Afflerbach, P., Beers, J., Boyd, C.D., Blachowicz, C., & Diffily, D. (2000). Reading : Surprise Me / Edition 1 . Reading : Surprise Me / Edition 1. Boston: Addison-Wesley Educational Publishers, Inc.

Alloway, T.P., Gathercole, S.E., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development*, 80(2), 606–621.

Beech, J.R. (2004). Using a dictionary: its influence on children's reading, spelling, and phonology. *Reading Psychology*, 25,19–36.

Borst, I. (2010). Understanding crowdsourcing; Effects of motivation and rewards on participation and performance in online voluntary activities. PhD thesis Erasmus Research Institute of Management, Erasmus University.

Brabham, D. C. (2012). Crowdsourcing: A model for leveraging online communities. In A. Delwiche & J. J. Henderson (Eds.), *The Participatory Cultures Handbook. (pp.120-129). UK:*Routledge.

Denh, M, (2008) *Working memory and academic learning: assessment and intervention.* Hoboken, NJ: John Wiley & Sons.

De Schryver, G-M. 2003. Lexicographers' Dreams in the Electronic-Dictionary Age. *International Journal of Lexicography* 16(2): 143-199.

Fillmore, C.J. (2003). Double-Decker Definitions: The role of frames in meaning explanations Sign Language Studies, 3(3), Spring 2003, 263-295.

Hamilton, H. (2012). The efficacy of dictionary use while reading for learning new words. *American Annals of the Deaf,* 157(4), 358-372.

Johnson, S. (1755). *A Dictionary of the English Language.* London: J. & P. Knapton.

Lew, R. (2010). New ways of indicating meaning in electronic dictionaries: Hope or hype? In Y. Zhang, (Ed.), *Learner's Lexicography and Second Language Teaching.* (pp. 387-404.Shanghai: Shanghai Foreign Language Education Press.

Weld, D.S., Adar, E., Chilton, L., Hoffmann, R., Horowitz, E., Koch, M., Landay, J., Lin, C.H. & Mausam. (2012). Personalized Online Education -A Crowdsourcing Challenge. *Human Computation AAAI Technical Report*, WS-12-08.

Winchester, S. (1998). *The Professor and the Madman: A Tale of Murder, Insanity, and the Making of the Oxford English Dictionary*, New York: Harper Perennial.

.

_

Dictionary 3.0 (D3)-Budget Justification
Georgia Tech Personnel
Dr. Harley Hamilton *(funded at 1.5 months)* is a Research Scientist at Georgia Tech. He will be the PI and develop oversee all aspects of the project.

Scott Robertson-Programmers *(funded at 2 months)* will build all data infrastructure and apps.

Consultants

Dr. Don McCreary *(2 days at $1000/day)* will provide guidance in the development of the D3 lexicon.

Dr. Daren Brabham *(5 days at $1000/day)* will develop the crowdsourcing plan.

Materials and Supplies
The materials and supplies budget includes money for programming software.
Travel
This budget includes $500 for attending the NEH meeting.
Additional Costs
Computer charges are calculated at $4,950 per year per full-time employee (current rates – we anticipate a 5% increase in year 1 of the project and every year thereafter), and are prorated based on actual amount of time worked on the project.
Fringe benefits are charged on all Georgia Tech at 28.5% per year except for Dr. Hamilton (1.4%). Indirect rates are charged on all direct costs less equipment at a rate of 52.7%.