



## IBM Faculty Awards printable version

IBM Faculty Awards are intended to recognize outstanding achievement and encourage exploratory projects of interest to IBM.

All nominations for these awards must be submitted by university Faculty Members to be considered for an award.

**Note:** Once submitted you will be unable to return to your nomination. Saving this nomination will allow you to return later using your URL and password. To retain a copy of lengthy answers you may submit, it is recommended that you use a word processing application to prepare the text. You can then do a copy/paste into the appropriate field.

### Faculty Awards nominee information

Please tell us about yourself in this section

Nominee information provided will not be used for any purpose beyond the scope of the IBM Faculty Awards program.

**First name:**

Edward

**Middle initial:**

Alan

**Last name:**

Fox

**E-mail address:**

fox@vt.edu

**Home page URL:**

<http://fox.cs.vt.edu>

**Country:**

United States

**University:**

Virginia Polytechnic Institute and State University(United States)

**If other, please specify**

**University address:**

Department of Computer Science  
660 McBryde Hall, M/C 0106

Virginia Tech

**City:**  
Blacksburg

**State/Province:**  
VA

**ZIP code/postal code:**  
24061

**Phone:**  
+1-540-231-5113

**Fax:**  
+1-540-231-6075

**Department:**  
Computer Science

**Nominee's curriculum vitae or resume**

Please see very long resume at  
<http://fox.cs.vt.edu/cv.htm>

**EDUCATION:**

8/83 Ph.D. Computer Science, Cornell University  
Area: information storage and retrieval  
Minor: linguistics; Advisor: Gerard Salton  
Title: Extending the Boolean and Vector Space  
Models of Information Retrieval with P-Norm  
Queries and Multiple Concept Types  
1/81 M.S. Computer Science, Cornell University  
2/72 B.S. Electrical Engineering (Computer Science  
Option), M.I.T.

**PROFESSIONAL SERVICE:**

**CURRENT - major responsibilities:**

2004- Chairman, IEEE-CS Technical Committee  
on Digital Libraries

2001- Director, Computing and Information  
Technology Interactive Digital Educational  
Library (CITIDEL), [www.citidel.org](http://www.citidel.org)

1996- Executive Director and Founder, Networked  
Digital Library of Theses and Dissertations  
(NDLTD, incorporated 5/2003),  
<http://www.ndltd.org>

**CURRENT - Selected Editor/Editorial Board Member:**

2004- Editorial board, International Journal on  
Digital Libraries ([www.dljournal.org](http://www.dljournal.org))

2002- Editorial board, ACM Transactions on  
Information Systems

2000- Co-editor-in-chief, ACM Journal of Education  
Resources In Computing (JERIC - and lead  
proposer 1998-2000),  
<http://www.acm.org/pubs/jeric/>

1995- Editor, Morgan Kaufmann Publishers, Inc.  
 Series on Multimedia Information and  
 Systems  
 1994- Editorial board, Multimedia Tools and  
 Applications (Kluwer)  
 1994- Foundation Editor, The Journal for Universal  
 Computer Science (Springer)  
 1990- Editorial board, Journal of Educational  
 Multimedia and Hypermedia (AAACE)  
 1987- Editorial board, Information Processing &  
 Management (Elsevier)  
 CURRENT - conferences/workshops (selected):  
 Member, Steering Committee, Joint Conference on  
 Digital Libraries (JCDL)  
 Member, Steering Committee, International Conf.  
 of Asian Digital Libraries (ICADL)  
 PC chair of the Information Retrieval track,  
 CIKM 2006  
 Member, JCDL 2006 program committee  
 Area Coordinator (AC) for SIGIR 2006, for Web  
 IR and Digital Libraries  
 Member, Program Committee, SPIRE 2006  
 CURRENT - other (selected):  
 2005- Member, Computational and Informational  
 Sciences Directorate Review Committee,  
 Pacific Northwest National Laboratory (PNNL) 2005- Member, Advisory

Group for Programmes  
 on Digital Repositories and on Digital Asset  
 Management and Preservation (assisting  
 JISC, the NSF-equivalent in UK)  
 2004- Member, Advisory Board, DELOS Network  
 of Excellence on Digital Libraries  
[\(http://www.delos.info/\)](http://www.delos.info/)  
 CONFERENCES/WORKSHOPS CHAIRED (selected):  
 Chair: Indo-US Workshop on Open Digital  
 Libraries and Interoperability, Arlington, VA,  
 June 23-25, 2003  
 Chair: First ACM/IEEE-CS Joint Conference on  
 Digital Libraries, JCDL'2001, Hotel Roanoke  
 and Conference Center, Roanoke, VA,  
 June 24-28, 2001  
 Program Chair: ACM Digital Libraries '99,  
 Berkeley, CA, Aug. 11-14, 1999  
 Chair: IBM Renaissance Meeting, Oct. 20-22,  
 1997, Virginia Tech, Blacksburg, VA  
 Program Chair: Digital Libraries'96, First ACM  
 International Conference on Digital Libraries,  
 March 20-23, 1996, Bethesda, MD  
 Program chair: 18th Intern'l Conf. on R & D in  
 Information Retrieval, SIGIR '95, July, 1995,  
 Seattle, WA

#### PUBLICATIONS:

#### VIDEOTAPE:

E. Fox, executive producer, script writer.

Interactive Digital Video. Five minute preview (Aug. 1989), hour final program (Jan. 1990), supplementing July 1989 Communications of the ACM, produced for ACM, with assistance from IBM T. J. Watson Research Center (with producer/director Guillermo Pulido).  
 BOOKS: 12 (co)authored or (co)edited  
 JOURNAL/MAGAZINE ARTICLES: over 75  
 BOOK CHAPTERS: over 35  
 REFEREED CONFERENCE/WORKSHOP PAPERS: over 110  
 REFEREED POSTERS: over 10  
 UNREFEREED INVITED CONFERENCE/WORKSHOP PAPERS: over 35  
 AWARDS SERVING AS (CO)PRINCIPAL INVESTIGATOR: over 90  
 KEYNOTE/BANQUET/INT'L INVITED/DISTINGUISHED SPEAKER PRESENTATIONS: over 50  
 TUTORIALS: over 65 in over 23 countries  
 DEMONSTRATIONS: over 20  
 ORAL PRESENTATIONS: over 245  
 TECHNICAL REPORTS, OTHER PUBLICATIONS: over 95  
 THESIS AND DISSERTATION ONLINE DOCUMENTS (Selected): over 20

### **Department head information**

Please tell us about your department head in this section

**First name:**

Dennis

**Last name:**

Kafura

**E-mail address:**

kafura@cs.vt.edu

### **Faculty Awards nominee technical background and interests**

Please tell us about the faculty nominee's technical background and interests in this section

**Select the main discipline that best describes this nomination**

Computer Science

**If Other main discipline, please specify**

**Select a sub-discipline**

Data Management

**If Other sub-discipline, please specify****Select additional sub-disciplines.**

Artificial Intelligence;Human-Computer Interaction;Knowledge Discovery and Data Mining;Natural Language Processing;Web

**Select a primary cross-disciplinary area of interest if applicable**

Human Resources

**If Other cross-discipline, please specify**

Please specify below up to three areas of technical interest (e.g., XML, Java, VLSI, etc.)

**Interest 1**

digital libraries

**Interest 2**

computing education

**Interest 3**

web services

**Questions****Award designation:**

Faculty member appointed more than 5 years of receiving his/her Ph.D.

**Project title**

Incorporating Superimposed Information into UIMA, and UIMA into Digital Library

Curriculum Development, as a Prototype Tool for  
Resource Collection Building and Automatic Syllabus Generation in the Semantic  
Web

**Description of project**

We 1) will leverage current work on at least 4 NSF funded projects to enhance UIMA, 2) will demonstrate and develop educational resources regarding the use of UIMA in support of the building of collections / digital libraries (DLs), and 3) will develop Web services to support education (first in computing, but applicable to other disciplines) through online annotated syllabi collections.

In the first project, DUE-0435059, "Collaborative Project: Superimposed Tools for Active Arrangement and Elaboration of Educational Resources," (<http://si.dlib.vt.edu/>) begun late in 2004, involving Portland State U. and Villanova U., we have explored technologies to facilitate handling of "marks", i.e., descriptions of parts of texts or multimedia documents, so that users can have middleware support for working at the level of sub-document granularity. In other words, these annotations of any works can be referred to, manipulated as first class objects, and resolved (bringing back the "base level" document with the annotations highlighted). Marks can refer to annotations created by users, or entities extracted through NLP, or sections of documents (e.g., a

reading list) identified by text analysis, or descriptions of popular topics identified by text mining. We plan to integrate the superimposed information technology into UIMA, to enrich and expand the ability of users and programs to refer to results of processing that has been aided by UIMA.

In the second related project, Virginia Tech's Department of Computer Science and UNC's School of Information and Library Science have been conducting an NSF-funded collaborative research project, IIS-0535057, "Curriculum Development: Digital Libraries"

(<http://curric.dlib.vt.edu/DLcurric.html>, PI Fox), since January 1, 2006. At least 18 digital library (DL) modules will be developed. Each module will include topics, lesson plans, suggested readings, textbooks, exercises, concept maps, and multimedia resource. Both one-semester and two-semester courses will be designed, tested and evaluated based on those modules. One of our initial efforts is to analyze syllabi from DL-related courses in the U.S. institutions and European institutions. From the syllabi, topics,

exercises, and text and multimedia resources need to be retrieved, semantically summarized, and classified against nine DL core modules. By working with the UIMA framework, we expect to automate the above process, populating each module with up-to-date, organized sets of heterogeneous resources, and to provide services

such as resource downloading and automatic syllabus generation for educators. We also will develop Web services for the Semantic Web that build upon syllabi, their various parts (described by marks as per above), and resources they refer to.

In the third related project, DUE-0532825, "Personalization of Content: Bridging the gap between NSDL and its users through the course website," led by Virginia Tech, but also involving Villanova U., begun late in 2005, we are working to help connect users with the resources of NSDL ([www.nsdlib.org](http://www.nsdlib.org), National Science Digital Library) by personalizing tools and systems they use. One part involves personalizing course syllabi, so each student's view is enriched in a way best for them. Other parts involve personalizing Courseware Management Systems like Sakai, Blackboard, and WebCT. In particular, when students refer to a particular part of the content for a course, they will see annotations and other enhancements to tailor and expand their learning.

In the fourth project, DUE-0333531, "Collaborative Project: The OCKHAM Library Network, Integrating the NSDL into Traditional Library Services," launched late in 2003, we have gained experience in developing Web services to support teaching and learning. We will leverage this work as we integrate our research with UIMA to prepare Web services that can enhance the Semantic Web.

We will involve the over 20 students in Virginia Tech's Digital Library Research Laboratory ([www.dlib.vt.edu](http://www.dlib.vt.edu)), and students (engaged in term projects) in courses taught by the PI: CS4624 (Multimedia, Hypertext, and Information Access) and CS5604 (Information Retrieval).

## REFERENCES

-----

1. Murthy, U., Torres, R.d.S. and Fox, E.A., A Superimposed Application for Enhanced Image Description and Retrieval. Submitted to the European conference on digital libraries (Alicante, Spain, 2006).

2. Manas Tungare, Xiaoyan Yu, Manuel Perez-Quinones, Edward A. Fox, Weiguo Fan, and Lillian Cassel, "Towards a Standardized Representation of Syllabi to Facilitate Sharing and Personalization of Digital Library Content". Submitted to the SW-EL workshop 2006.
3. Murthy, U., Ahuja, K., Murthy, S. and Fox, E.A., SIMPEL: A Superimposed Multimedia Presentation Editor and Player. Demo proposal to appear in proceedings of the Joint Conference on Digital Libraries 2006 (Chapel Hill, NC, USA).
4. SPARCE website, Datalab, Department of Computer Science, Portland State University, 2005, <http://datalab.cs.pdx.edu/sparce/>.
5. Fox, E.A. and Murthy, U. The Superimposed Information Project at Virginia Tech, 2005, <http://si.dlib.vt.edu/>.
6. Murthy, U. and Ahuja, K. SIMPEL: A SuperImposed Multimedia Presentation Editor and PLayer, Project report, Virginia Tech, Blacksburg, VA., 2005, <http://pubs.dlib.vt.edu:9090/48/>.
7. Baoping Zhang, Yuxin Chen, Weiguo Fan, Edward A. Fox, Marcos Andre Goncalves, Marco Cristo, Pável Calado. Intelligent fusion of structural and citation-based evidence for text classification. In Proc. SIGIR 2005, 28th Annual International ACM SIGIR Conference, Salvador, Brazil, August 15-19, 667-668, <http://doi.acm.org/10.1145/1076034.1076181>
8. S. Perugini, K. McDevitt, R. Richardson, M. Perez-Quinones, R. Shen, N. Ramakrishnan, C. Williams, and E. A. Fox. Enhancing Usability in CITIDEL: Multimodal, Multilingual, and Interactive Visualization Interfaces. In Proceedings of the Fourth ACM/IEEE Joint Conference on Digital Libraries: Global Reach and Diverse Impact, JCDL2004, Tucson, AZ, June 7-11, 315-324.
9. Murthy, S., Maier, D., Delcambre, L. and Bowers, S., Putting Integrated Information into Context: Superimposing Conceptual Models with SPARCE. In Proceedings of the First Asia-Pacific Conference of Conceptual Modeling (Denedin, New Zealand, 2004), 71-80.
10. E.A. Fox, Case Studies in the US National Science Digital Library: DL-in-a-Box, CITIDEL, and OCKHAM. Invited talk in: Tengku Mohd Tengku Sembok, Halimah Badioze Zaman, Hsinchun Chen, Shalini R. Urs, Sung Hyon Myaeng, eds. Digital Libraries: Technology and Management of Indigenous Knowledge for Global Access; Proceedings 6th International Conference on Asian Digital Libraries, ICADL 2003, Kuala Lumpur, Malaysia, Dec. 2003. Springer-Verlag GmbH, Lecture Notes in Computer Science, vol. 2911, 17-25
11. Hussein Suleman, Edward A Fox, Rohit Kelapure, Aaron Krowne, Ming Luo. Building digital libraries from simple building blocks, Online Information Review Volume 27 Number 5, Oct. 2003, 301-310
12. Maier, D. and Delcambre, L., Superimposed Information for the Internet. In Proceedings of the WebDB Workshop (1999), 1-9.

### **Project objectives and goals**

Our goals are 1) to enhance UIMA, 2) to expand the benefits of superimposed

information middleware, 3) to prototype UIMA-based support of analyzing Web pages related to course syllabi, 4) to turn that prototyping effort into educational resources for those studying about digital libraries (or related topics), 5) to prepare a rich collection of enhanced and annotated syllabi (about computing, but extensible to other fields), and 6) to develop Web services that will support the Semantic Web

First, we will leverage our work with superimposed information to enhance UIMA to support marks, and so that UIMA-based processing can lead to enriched output (including marks).

Second, we will leverage the work on personalizing education by utilizing UIMA's support for a Collection Processing Architecture, resource collections can be developed from the DL syllabus collection we are developing, by extracting reference paper titles from each syllabus, downloading and analyzing those papers, and populating a database with them for each DL core module. We will build a syllabus analysis component using UIMA, with the input of our collection of syllabi, and the output a structured document of a syllabus. The results of the analysis phase would be annotated syllabi. The deliverable would be Web services for semantic search/browse on syllabi, where users can operate on the level of whole syllabi, as well as on the level of particular interesting components. We will standardize a syllabus XML schema similar to what RSS did with news feed. In effect an ontology of syllabi would be built. Advanced value-adding Web services, such as query, browse, and automatic generation of syllabi, will be provided.

Third, we will connect that work with our digital library curriculum project, since DL syllabi are of particular interest.

Fourth, we will extend this work into Web services to enhance the Semantic Web.

Our objectives include:

- a) integrating our superimposed information research work with UIMA,
- b) integrating our DL curriculum work with UIMA so that courses on DLs (and on related topics like information retrieval) are enhanced,
- c) integrating our work with NSDL and personalization so that use of UIMA can help large numbers of those who can benefit from NSDL,
- d) integrating our work with OCKHAM so that all of the above lead to Web services that enhance the Semantic Web.

### **Long term impact to the information/computing industry and IBM**

The people who are involved in DL curriculum development and DL courses -- including students, instructors, expert advisory group members, and the ones developing curriculum for other areas -- will become familiar with UIMA and its capabilities. This will extend to an even larger group interested in any of the modules under development, including those studying information retrieval.

Our work will promote the use of UIMA among those groups, and in academia in general, by enhancing the ideas for component development and new applicable areas. These new ideas will contribute to the advancement of information/computing industry considering that UIMA and its components are freely available and many companies are willing to use them. Further, through integrating support for superimposed information (SI), we will expand the set of people interested in that to include those who will benefit from enhanced annotation

and other SI-related capabilities.

Anyone who uses a syllabus can ultimately benefit, ultimately every teach and learner, since we plan to enrich the genre, to use analysis to normalize current syllabi, and to provide generic Web services that help people and systems (in the Semantic Web) to benefit from enhanced syllabi.

**Amount of cash requested, in US Dollars. This amount cannot exceed \$40000.**  
\$ 40000

**Planned use for cash (The IBM Faculty Award is a gift and is processed as a donation to your university. We expect that no overhead will be charged against this award.)**

\* Fund graduate research assistants

\* Fund PI for 2 weeks

**If awarded, a check will be made payable to the university named in this nomination. If this is not the case please specify a different name here. NOTE: All US award recipients will be required to complete a copy of IRS form 501(c)(3) that will be provided by IBM at a later time.**

**If awarded, will the award check be made payable to a non-US school?**

No

#### **Data privacy statement**

According to the Privacy Law we are required to inform you on the processing of your personal data and on your rights as indicated in this Privacy Statement. This data may be used by IBM or selected organisations to provide you with information about other offerings for the purposes of direct marketing. To receive this via any means including automated calling systems without human intervention (e.g. e-mail, facsimile, MMS - or SMS - type messages or other electronic means) check the first box below. Alternatively, if you would prefer not to receive such information by any means, check the second box.

May use information

Do not send email

Correspondence via e-mail is necessary to complete this nomination process. Submission of this nomination implies permission to communicate with all participants via e-mail in matters related to this program.