

**02 INFORMATION ABOUT PRINCIPAL INVESTIGATORS/PROJECT DIRECTORS(PI/PD) and  
co-PRINCIPAL INVESTIGATORS/co-PROJECT DIRECTORS**

---

Submit only ONE copy of this form for each PI/PD and co-PI/PD identified on the proposal. The form(s) should be attached to the original proposal as specified in GPG Section II.C.a. Submission of this information is voluntary and is not a precondition of award. This information will not be disclosed to external peer reviewers. **DO NOT INCLUDE THIS FORM WITH ANY OF THE OTHER COPIES OF YOUR PROPOSAL AS THIS MAY COMPROMISE THE CONFIDENTIALITY OF THE INFORMATION.**

---

**PI/PD Name:** Michael J Khoo

**Gender:**  Male  Female  
**Ethnicity:** (Choose one response)  Hispanic or Latino  Not Hispanic or Latino

**Race:**  
(Select one or more)  American Indian or Alaska Native  
 Asian  
 Black or African American  
 Native Hawaiian or Other Pacific Islander  
 White

**Disability Status:**  
(Select one or more)  Hearing Impairment  
 Visual Impairment  
 Mobility/Orthopedic Impairment  
 Other  
 None

**Citizenship:** (Choose one)  U.S. Citizen  Permanent Resident  Other non-U.S. Citizen

**Check here if you do not wish to provide any or all of the above information (excluding PI/PD name):**

**REQUIRED: Check here if you are currently serving (or have previously served) as a PI, co-PI or PD on any federally funded project**

---

**Ethnicity Definition:**

**Hispanic or Latino.** A person of Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race.

**Race Definitions:**

**American Indian or Alaska Native.** A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.

**Asian.** A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.

**Black or African American.** A person having origins in any of the black racial groups of Africa.

**Native Hawaiian or Other Pacific Islander.** A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.

**White.** A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

---

**WHY THIS INFORMATION IS BEING REQUESTED:**

The Federal Government has a continuing commitment to monitor the operation of its review and award processes to identify and address any inequities based on gender, race, ethnicity, or disability of its proposed PIs/PDs. To gather information needed for this important task, the proposer should submit a single copy of this form for each identified PI/PD with each proposal. Submission of the requested information is voluntary and will not affect the organization's eligibility for an award. However, information not submitted will seriously undermine the statistical validity, and therefore the usefulness, of information received from others. Any individual not wishing to submit some or all the information should check the box provided for this purpose. (The exceptions are the PI/PD name and the information about prior Federal support, the last question above.)

Collection of this information is authorized by the NSF Act of 1950, as amended, 42 U.S.C. 1861, et seq. Demographic data allows NSF to gauge whether our programs and other opportunities in science and technology are fairly reaching and benefiting everyone regardless of demographic category; to ensure that those in under-represented groups have the same knowledge of and access to programs and other research and educational opportunities; and to assess involvement of international investigators in work supported by NSF. The information may be disclosed to government contractors, experts, volunteers and researchers to complete assigned work; and to other government agencies in order to coordinate and assess programs. The information may be added to the Reviewer file and used to select potential candidates to serve as peer reviewers or advisory committee members. See Systems of Records, NSF-50, "Principal Investigator/Proposal File and Associated Records", 63 Federal Register 267 (January 5, 1998), and NSF-51, "Reviewer/Proposal File and Associated Records", 63 Federal Register 268 (January 5, 1998).

**02 INFORMATION ABOUT PRINCIPAL INVESTIGATORS/PROJECT DIRECTORS(PI/PD) and  
co-PRINCIPAL INVESTIGATORS/co-PROJECT DIRECTORS**

---

Submit only ONE copy of this form for each PI/PD and co-PI/PD identified on the proposal. The form(s) should be attached to the original proposal as specified in GPG Section II.C.a. Submission of this information is voluntary and is not a precondition of award. This information will not be disclosed to external peer reviewers. **DO NOT INCLUDE THIS FORM WITH ANY OF THE OTHER COPIES OF YOUR PROPOSAL AS THIS MAY COMPROMISE THE CONFIDENTIALITY OF THE INFORMATION.**

---

**PI/PD Name:** Eileen Abels

**Gender:**  Male  Female  
**Ethnicity:** (Choose one response)  Hispanic or Latino  Not Hispanic or Latino

**Race:**  
(Select one or more)  
 American Indian or Alaska Native  
 Asian  
 Black or African American  
 Native Hawaiian or Other Pacific Islander  
 White

**Disability Status:**  
(Select one or more)  
 Hearing Impairment  
 Visual Impairment  
 Mobility/Orthopedic Impairment  
 Other  
 None

**Citizenship:** (Choose one)  U.S. Citizen  Permanent Resident  Other non-U.S. Citizen

**Check here if you do not wish to provide any or all of the above information (excluding PI/PD name):**

**REQUIRED: Check here if you are currently serving (or have previously served) as a PI, co-PI or PD on any federally funded project**

---

**Ethnicity Definition:**

**Hispanic or Latino.** A person of Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race.

**Race Definitions:**

**American Indian or Alaska Native.** A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.

**Asian.** A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.

**Black or African American.** A person having origins in any of the black racial groups of Africa.

**Native Hawaiian or Other Pacific Islander.** A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.

**White.** A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

---

**WHY THIS INFORMATION IS BEING REQUESTED:**

The Federal Government has a continuing commitment to monitor the operation of its review and award processes to identify and address any inequities based on gender, race, ethnicity, or disability of its proposed PIs/PDs. To gather information needed for this important task, the proposer should submit a single copy of this form for each identified PI/PD with each proposal. Submission of the requested information is voluntary and will not affect the organization's eligibility for an award. However, information not submitted will seriously undermine the statistical validity, and therefore the usefulness, of information received from others. Any individual not wishing to submit some or all the information should check the box provided for this purpose. (The exceptions are the PI/PD name and the information about prior Federal support, the last question above.)

Collection of this information is authorized by the NSF Act of 1950, as amended, 42 U.S.C. 1861, et seq. Demographic data allows NSF to gauge whether our programs and other opportunities in science and technology are fairly reaching and benefiting everyone regardless of demographic category; to ensure that those in under-represented groups have the same knowledge of and access to programs and other research and educational opportunities; and to assess involvement of international investigators in work supported by NSF. The information may be disclosed to government contractors, experts, volunteers and researchers to complete assigned work; and to other government agencies in order to coordinate and assess programs. The information may be added to the Reviewer file and used to select potential candidates to serve as peer reviewers or advisory committee members. See Systems of Records, NSF-50, "Principal Investigator/Proposal File and Associated Records", 63 Federal Register 267 (January 5, 1998), and NSF-51, "Reviewer/Proposal File and Associated Records", 63 Federal Register 268 (January 5, 1998).

**02 INFORMATION ABOUT PRINCIPAL INVESTIGATORS/PROJECT DIRECTORS(PI/PD) and  
co-PRINCIPAL INVESTIGATORS/co-PROJECT DIRECTORS**

---

Submit only ONE copy of this form for each PI/PD and co-PI/PD identified on the proposal. The form(s) should be attached to the original proposal as specified in GPG Section II.C.a. Submission of this information is voluntary and is not a precondition of award. This information will not be disclosed to external peer reviewers. **DO NOT INCLUDE THIS FORM WITH ANY OF THE OTHER COPIES OF YOUR PROPOSAL AS THIS MAY COMPROMISE THE CONFIDENTIALITY OF THE INFORMATION.**

---

**PI/PD Name:** Sean Goggins

**Gender:**  Male  Female  
**Ethnicity:** (Choose one response)  Hispanic or Latino  Not Hispanic or Latino

**Race:**  
(Select one or more)  
 American Indian or Alaska Native  
 Asian  
 Black or African American  
 Native Hawaiian or Other Pacific Islander  
 White

**Disability Status:**  
(Select one or more)  
 Hearing Impairment  
 Visual Impairment  
 Mobility/Orthopedic Impairment  
 Other  
 None

**Citizenship:** (Choose one)  U.S. Citizen  Permanent Resident  Other non-U.S. Citizen

**Check here if you do not wish to provide any or all of the above information (excluding PI/PD name):**

**REQUIRED: Check here if you are currently serving (or have previously served) as a PI, co-PI or PD on any federally funded project**

---

**Ethnicity Definition:**

**Hispanic or Latino.** A person of Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race.

**Race Definitions:**

**American Indian or Alaska Native.** A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.

**Asian.** A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.

**Black or African American.** A person having origins in any of the black racial groups of Africa.

**Native Hawaiian or Other Pacific Islander.** A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.

**White.** A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

---

**WHY THIS INFORMATION IS BEING REQUESTED:**

The Federal Government has a continuing commitment to monitor the operation of its review and award processes to identify and address any inequities based on gender, race, ethnicity, or disability of its proposed PIs/PDs. To gather information needed for this important task, the proposer should submit a single copy of this form for each identified PI/PD with each proposal. Submission of the requested information is voluntary and will not affect the organization's eligibility for an award. However, information not submitted will seriously undermine the statistical validity, and therefore the usefulness, of information received from others. Any individual not wishing to submit some or all the information should check the box provided for this purpose. (The exceptions are the PI/PD name and the information about prior Federal support, the last question above.)

Collection of this information is authorized by the NSF Act of 1950, as amended, 42 U.S.C. 1861, et seq. Demographic data allows NSF to gauge whether our programs and other opportunities in science and technology are fairly reaching and benefiting everyone regardless of demographic category; to ensure that those in under-represented groups have the same knowledge of and access to programs and other research and educational opportunities; and to assess involvement of international investigators in work supported by NSF. The information may be disclosed to government contractors, experts, volunteers and researchers to complete assigned work; and to other government agencies in order to coordinate and assess programs. The information may be added to the Reviewer file and used to select potential candidates to serve as peer reviewers or advisory committee members. See Systems of Records, NSF-50, "Principal Investigator/Proposal File and Associated Records", 63 Federal Register 267 (January 5, 1998), and NSF-51, "Reviewer/Proposal File and Associated Records", 63 Federal Register 268 (January 5, 1998).

**02 INFORMATION ABOUT PRINCIPAL INVESTIGATORS/PROJECT DIRECTORS(PI/PD) and  
co-PRINCIPAL INVESTIGATORS/co-PROJECT DIRECTORS**

---

Submit only ONE copy of this form for each PI/PD and co-PI/PD identified on the proposal. The form(s) should be attached to the original proposal as specified in GPG Section II.C.a. Submission of this information is voluntary and is not a precondition of award. This information will not be disclosed to external peer reviewers. **DO NOT INCLUDE THIS FORM WITH ANY OF THE OTHER COPIES OF YOUR PROPOSAL AS THIS MAY COMPROMISE THE CONFIDENTIALITY OF THE INFORMATION.**

---

**PI/PD Name:**           Jiexun Li          

**Gender:**  Male  Female  
**Ethnicity:** (Choose one response)  Hispanic or Latino  Not Hispanic or Latino

**Race:**  
(Select one or more)  American Indian or Alaska Native  
 Asian  
 Black or African American  
 Native Hawaiian or Other Pacific Islander  
 White

**Disability Status:**  
(Select one or more)  Hearing Impairment  
 Visual Impairment  
 Mobility/Orthopedic Impairment  
 Other  
 None

**Citizenship:** (Choose one)  U.S. Citizen  Permanent Resident  Other non-U.S. Citizen

**Check here if you do not wish to provide any or all of the above information (excluding PI/PD name):**

**REQUIRED: Check here if you are currently serving (or have previously served) as a PI, co-PI or PD on any federally funded project**

---

**Ethnicity Definition:**

**Hispanic or Latino.** A person of Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race.

**Race Definitions:**

**American Indian or Alaska Native.** A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.

**Asian.** A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.

**Black or African American.** A person having origins in any of the black racial groups of Africa.

**Native Hawaiian or Other Pacific Islander.** A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.

**White.** A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

---

**WHY THIS INFORMATION IS BEING REQUESTED:**

The Federal Government has a continuing commitment to monitor the operation of its review and award processes to identify and address any inequities based on gender, race, ethnicity, or disability of its proposed PIs/PDs. To gather information needed for this important task, the proposer should submit a single copy of this form for each identified PI/PD with each proposal. Submission of the requested information is voluntary and will not affect the organization's eligibility for an award. However, information not submitted will seriously undermine the statistical validity, and therefore the usefulness, of information received from others. Any individual not wishing to submit some or all the information should check the box provided for this purpose. (The exceptions are the PI/PD name and the information about prior Federal support, the last question above.)

Collection of this information is authorized by the NSF Act of 1950, as amended, 42 U.S.C. 1861, et seq. Demographic data allows NSF to gauge whether our programs and other opportunities in science and technology are fairly reaching and benefiting everyone regardless of demographic category; to ensure that those in under-represented groups have the same knowledge of and access to programs and other research and educational opportunities; and to assess involvement of international investigators in work supported by NSF. The information may be disclosed to government contractors, experts, volunteers and researchers to complete assigned work; and to other government agencies in order to coordinate and assess programs. The information may be added to the Reviewer file and used to select potential candidates to serve as peer reviewers or advisory committee members. See Systems of Records, NSF-50, "Principal Investigator/Proposal File and Associated Records", 63 Federal Register 267 (January 5, 1998), and NSF-51, "Reviewer/Proposal File and Associated Records", 63 Federal Register 268 (January 5, 1998).

**02 INFORMATION ABOUT PRINCIPAL INVESTIGATORS/PROJECT DIRECTORS(PI/PD) and  
co-PRINCIPAL INVESTIGATORS/co-PROJECT DIRECTORS**

---

Submit only ONE copy of this form for each PI/PD and co-PI/PD identified on the proposal. The form(s) should be attached to the original proposal as specified in GPG Section II.C.a. Submission of this information is voluntary and is not a precondition of award. This information will not be disclosed to external peer reviewers. **DO NOT INCLUDE THIS FORM WITH ANY OF THE OTHER COPIES OF YOUR PROPOSAL AS THIS MAY COMPROMISE THE CONFIDENTIALITY OF THE INFORMATION.**

---

**PI/PD Name:** Gerry Stahl

**Gender:**  Male  Female  
**Ethnicity:** (Choose one response)  Hispanic or Latino  Not Hispanic or Latino

**Race:**  
(Select one or more)  
 American Indian or Alaska Native  
 Asian  
 Black or African American  
 Native Hawaiian or Other Pacific Islander  
 White

**Disability Status:**  
(Select one or more)  
 Hearing Impairment  
 Visual Impairment  
 Mobility/Orthopedic Impairment  
 Other  
 None

**Citizenship:** (Choose one)  U.S. Citizen  Permanent Resident  Other non-U.S. Citizen

**Check here if you do not wish to provide any or all of the above information (excluding PI/PD name):**

**REQUIRED: Check here if you are currently serving (or have previously served) as a PI, co-PI or PD on any federally funded project**

---

**Ethnicity Definition:**

**Hispanic or Latino.** A person of Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race.

**Race Definitions:**

**American Indian or Alaska Native.** A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.

**Asian.** A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.

**Black or African American.** A person having origins in any of the black racial groups of Africa.

**Native Hawaiian or Other Pacific Islander.** A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.

**White.** A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

---

**WHY THIS INFORMATION IS BEING REQUESTED:**

The Federal Government has a continuing commitment to monitor the operation of its review and award processes to identify and address any inequities based on gender, race, ethnicity, or disability of its proposed PIs/PDs. To gather information needed for this important task, the proposer should submit a single copy of this form for each identified PI/PD with each proposal. Submission of the requested information is voluntary and will not affect the organization's eligibility for an award. However, information not submitted will seriously undermine the statistical validity, and therefore the usefulness, of information received from others. Any individual not wishing to submit some or all the information should check the box provided for this purpose. (The exceptions are the PI/PD name and the information about prior Federal support, the last question above.)

Collection of this information is authorized by the NSF Act of 1950, as amended, 42 U.S.C. 1861, et seq. Demographic data allows NSF to gauge whether our programs and other opportunities in science and technology are fairly reaching and benefiting everyone regardless of demographic category; to ensure that those in under-represented groups have the same knowledge of and access to programs and other research and educational opportunities; and to assess involvement of international investigators in work supported by NSF. The information may be disclosed to government contractors, experts, volunteers and researchers to complete assigned work; and to other government agencies in order to coordinate and assess programs. The information may be added to the Reviewer file and used to select potential candidates to serve as peer reviewers or advisory committee members. See Systems of Records, NSF-50, "Principal Investigator/Proposal File and Associated Records", 63 Federal Register 267 (January 5, 1998), and NSF-51, "Reviewer/Proposal File and Associated Records", 63 Federal Register 268 (January 5, 1998).

## List of Suggested Reviewers or Reviewers Not To Include (optional)

---

### **SUGGESTED REVIEWERS:**

Not Listed

### **REVIEWERS NOT TO INCLUDE:**

Not Listed

---

## COVER SHEET FOR PROPOSAL TO THE NATIONAL SCIENCE FOUNDATION

PROGRAM ANNOUNCEMENT/SOLICITATION NO./CLOSING DATE/if not in response to a program announcement/solicitation enter NSF 09-29					<b>FOR NSF USE ONLY</b>	
NSF 09-559			09/21/09		<b>NSF PROPOSAL NUMBER</b>	
FOR CONSIDERATION BY NSF ORGANIZATION UNIT(S) (Indicate the most specific unit known, i.e. program, division, etc.)					<b>0968542</b>	
<b>SES - SCIENCE, TECH &amp; SOCIETY</b>						
DATE RECEIVED	NUMBER OF COPIES	DIVISION ASSIGNED	FUND CODE	DUNS# (Data Universal Numbering System)	FILE LOCATION	
09/21/2009	2	04050000 SES	7603	002604817	09/22/2009 9:11am	
EMPLOYER IDENTIFICATION NUMBER (EIN) OR TAXPAYER IDENTIFICATION NUMBER (TIN)		SHOW PREVIOUS AWARD NO. IF THIS IS <input type="checkbox"/> A RENEWAL <input type="checkbox"/> AN ACCOMPLISHMENT-BASED RENEWAL		IS THIS PROPOSAL BEING SUBMITTED TO ANOTHER FEDERAL AGENCY? YES <input type="checkbox"/> NO <input checked="" type="checkbox"/> IF YES, LIST ACRONYM(S)		
231352630						
NAME OF ORGANIZATION TO WHICH AWARD SHOULD BE MADE			ADDRESS OF AWARDEE ORGANIZATION, INCLUDING 9 DIGIT ZIP CODE			
Drexel University			Drexel University 3201 Arch Street Philadelphia, PA. 191042737			
AWARDEE ORGANIZATION CODE (IF KNOWN)						
0032565000						
NAME OF PERFORMING ORGANIZATION, IF DIFFERENT FROM ABOVE			ADDRESS OF PERFORMING ORGANIZATION, IF DIFFERENT, INCLUDING 9 DIGIT ZIP CODE			
PERFORMING ORGANIZATION CODE (IF KNOWN)						
IS AWARDEE ORGANIZATION (Check All That Apply) (See GPG II.C For Definitions)		<input type="checkbox"/> SMALL BUSINESS <input type="checkbox"/> FOR-PROFIT ORGANIZATION		<input type="checkbox"/> MINORITY BUSINESS <input type="checkbox"/> WOMAN-OWNED BUSINESS		<input type="checkbox"/> IF THIS IS A PRELIMINARY PROPOSAL THEN CHECK HERE
TITLE OF PROPOSED PROJECT <b>Collaborative Knowledge Work in Socially Intelligent Computational Systems</b>						
REQUESTED AMOUNT \$	PROPOSED DURATION (1-60 MONTHS)	REQUESTED STARTING DATE	SHOW RELATED PRELIMINARY PROPOSAL NO. IF APPLICABLE			
747,599	36 months	01/01/10				
CHECK APPROPRIATE BOX(ES) IF THIS PROPOSAL INCLUDES ANY OF THE ITEMS LISTED BELOW						
<input checked="" type="checkbox"/> BEGINNING INVESTIGATOR (GPG I.G.2)						
<input type="checkbox"/> HUMAN SUBJECTS (GPG II.D.7) Human Subjects Assurance Number _____						
<input type="checkbox"/> DISCLOSURE OF LOBBYING ACTIVITIES (GPG II.C.1.e) Exemption Subsection _____ or IRB App. Date _____						
<input type="checkbox"/> PROPRIETARY & PRIVILEGED INFORMATION (GPG I.D, II.C.1.d)						
<input type="checkbox"/> INTERNATIONAL COOPERATIVE ACTIVITIES: COUNTRY/COUNTRIES INVOLVED (GPG II.C.2.j)						
<input type="checkbox"/> HISTORIC PLACES (GPG II.C.2.j)						
<input type="checkbox"/> EAGER* (GPG II.D.2) <input type="checkbox"/> RAPID** (GPG II.D.1)						
<input type="checkbox"/> HIGH RESOLUTION GRAPHICS/OTHER GRAPHICS WHERE EXACT COLOR REPRESENTATION IS REQUIRED FOR PROPER INTERPRETATION (GPG I.G.1)						
<input type="checkbox"/> VERTEBRATE ANIMALS (GPG II.D.6) IACUC App. Date _____						
PHS Animal Welfare Assurance Number _____						
PI/PD DEPARTMENT			PI/PD POSTAL ADDRESS			
College of Information Science and Techn			3141 Chestnut Street			
PI/PD FAX NUMBER			Philadelphia, PA 19104			
215-895-2494			United States			
NAMES (TYPED)	High Degree	Yr of Degree	Telephone Number	Electronic Mail Address		
PI/PD NAME						
Michael J Khoo	PhD	2004	215-895-2474	mjkhoo@gmail.com		
CO-PI/PD						
Eileen Abels	PhD	1985	215-895-2000	Eileen.Abels@ischool.drexel.edu		
CO-PI/PD						
Sean Goggins	UKNW	2009	215-895-5849	sgoggins@drexel.edu		
CO-PI/PD						
Jiexun Li	PhD	2007	215-895-5849	jl622@drexel.edu		
CO-PI/PD						
Gerry Stahl	PhD	1993	215-895-0544	Gerry.Stahl@drexel.edu		

## CERTIFICATION PAGE

### Certification for Authorized Organizational Representative or Individual Applicant:

By signing and submitting this proposal, the Authorized Organizational Representative or Individual Applicant is: (1) certifying that statements made herein are true and complete to the best of his/her knowledge; and (2) agreeing to accept the obligation to comply with NSF award terms and conditions if an award is made as a result of this application. Further, the applicant is hereby providing certifications regarding debarment and suspension, drug-free workplace, and lobbying activities (see below), nondiscrimination, and flood hazard insurance (when applicable) as set forth in the NSF Proposal & Award Policies & Procedures Guide, Part I: the Grant Proposal Guide (GPG) (NSF 09-29). Willful provision of false information in this application and its supporting documents or in reports required under an ensuing award is a criminal offense (U. S. Code, Title 18, Section 1001).

### Conflict of Interest Certification

In addition, if the applicant institution employs more than fifty persons, by electronically signing the NSF Proposal Cover Sheet, the Authorized Organizational Representative of the applicant institution is certifying that the institution has implemented a written and enforced conflict of interest policy that is consistent with the provisions of the NSF Proposal & Award Policies & Procedures Guide, Part II, Award & Administration Guide (AAG) Chapter IV.A; that to the best of his/her knowledge, all financial disclosures required by that conflict of interest policy have been made; and that all identified conflicts of interest will have been satisfactorily managed, reduced or eliminated prior to the institution's expenditure of any funds under the award, in accordance with the institution's conflict of interest policy. Conflicts which cannot be satisfactorily managed, reduced or eliminated must be disclosed to NSF.

### Drug Free Work Place Certification

By electronically signing the NSF Proposal Cover Sheet, the Authorized Organizational Representative or Individual Applicant is providing the Drug Free Work Place Certification contained in Exhibit II-3 of the Grant Proposal Guide.

### Debarment and Suspension Certification

(If answer "yes", please provide explanation.)

Is the organization or its principals presently debarred, suspended, proposed for debarment, declared ineligible, or voluntarily excluded from covered transactions by any Federal department or agency?

Yes

No

By electronically signing the NSF Proposal Cover Sheet, the Authorized Organizational Representative or Individual Applicant is providing the Debarment and Suspension Certification contained in Exhibit II-4 of the Grant Proposal Guide.

### Certification Regarding Lobbying

The following certification is required for an award of a Federal contract, grant, or cooperative agreement exceeding \$100,000 and for an award of a Federal loan or a commitment providing for the United States to insure or guarantee a loan exceeding \$150,000.

### Certification for Contracts, Grants, Loans and Cooperative Agreements

The undersigned certifies, to the best of his or her knowledge and belief, that:

- (1) No federal appropriated funds have been paid or will be paid, by or on behalf of the undersigned, to any person for influencing or attempting to influence an officer or employee of any agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with the awarding of any federal contract, the making of any Federal grant, the making of any Federal loan, the entering into of any cooperative agreement, and the extension, continuation, renewal, amendment, or modification of any Federal contract, grant, loan, or cooperative agreement.
- (2) If any funds other than Federal appropriated funds have been paid or will be paid to any person for influencing or attempting to influence an officer or employee of any agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with this Federal contract, grant, loan, or cooperative agreement, the undersigned shall complete and submit Standard Form-LLL, "Disclosure of Lobbying Activities," in accordance with its instructions.
- (3) The undersigned shall require that the language of this certification be included in the award documents for all subawards at all tiers including subcontracts, subgrants, and contracts under grants, loans, and cooperative agreements and that all subrecipients shall certify and disclose accordingly.

This certification is a material representation of fact upon which reliance was placed when this transaction was made or entered into. Submission of this certification is a prerequisite for making or entering into this transaction imposed by section 1352, Title 31, U.S. Code. Any person who fails to file the required certification shall be subject to a civil penalty of not less than \$10,000 and not more than \$100,000 for each such failure.

### Certification Regarding Nondiscrimination

By electronically signing the NSF Proposal Cover Sheet, the Authorized Organizational Representative is providing the Certification Regarding Nondiscrimination contained in Exhibit II-6 of the Grant Proposal Guide.

### Certification Regarding Flood Hazard Insurance

Two sections of the National Flood Insurance Act of 1968 (42 USC §4012a and §4106) bar Federal agencies from giving financial assistance for acquisition or construction purposes in any area identified by the Federal Emergency Management Agency (FEMA) as having special flood hazards unless the:

- (1) community in which that area is located participates in the national flood insurance program; and
- (2) building (and any related equipment) is covered by adequate flood insurance.

By electronically signing the NSF Proposal Cover Sheet, the Authorized Organizational Representative or Individual Applicant located in FEMA-designated special flood hazard areas is certifying that adequate flood insurance has been or will be obtained in the following situations:

- (1) for NSF grants for the construction of a building or facility, regardless of the dollar amount of the grant; and
- (2) for other NSF Grants when more than \$25,000 has been budgeted in the proposal for repair, alteration or improvement (construction) of a building or facility.

AUTHORIZED ORGANIZATIONAL REPRESENTATIVE		SIGNATURE	DATE
NAME <b>Margaret Vigiolto</b>		<b>Electronic Signature</b>	<b>Sep 22 2009 8:44AM</b>
TELEPHONE NUMBER <b>215-895-2311</b>	ELECTRONIC MAIL ADDRESS <b>vigiolto@drexel.edu</b>	FAX NUMBER <b>215-895-6699</b>	

\* EAGER - EARly-concept Grants for Exploratory Research

\*\* RAPID - Grants for Rapid Response Research



## **Collaborative Knowledge Work In Socially Intelligent Computational Systems**

Computer-based information systems connect users instantaneously across time and space. They promise to support communication, work and social collaboration, and creativity and productivity. However, this promise can play out in emergent, complex, and unexpected ways, at societal, organizational, and individual levels.

From a science, technology and society (STS) perspective, information systems, like all technologies, are complex systems of people, technologies, and social, economic, historical, and other components. It is the iterative interactions between these multiple components that result in unexpected social and technological outcomes. Current and proposed plans for cyberinfrastructure and social-computational systems make this observation more relevant than ever. The analysis and modeling of these systems and processes will require a deep understanding of how human intelligence and knowledge are mediated in and amongst groups in complex distributed computational environments. Such an approach requires in turn a sophisticated understanding of what intelligence and knowledge are and how they manifest in social-computational systems.

**We propose to study the relationships between human, organizational, and computational elements in a prototypical socially-intelligent computational system (the Collaborative Environment for Creating Metadata - CECM), with the aim of understanding how such a system supports communities of practice, tacit knowledge, and legitimate peripheral participation. The study will provide support for further inquiries into the requirements, affordances, and capabilities of such systems.**

Through a detailed investigation of one specific case the study will provide support for further inquiries into the knowledge requirements, affordances, and capabilities of such systems. The study will integrate both social and computational components. The social component includes models of the behavior of individuals and groups when engaged in online collaborative work, and the social dimensions of knowledge sharing in online settings. The computational component includes the development of social network workspaces and automatic recommender tools to support knowledge work and metadata generation by non-experts. The integration is built on analyses of communication and knowledge construction in these tools at both individual and group levels, using a range of different analytical techniques, including text and data mining, content analysis, discourse analysis, user metrics and others. It will study the interactions between the social and technical components, and will address central social-computational research questions such as:

- How do communities of practice form in social-computational systems?
- How does legitimate peripheral participation occur in social-computational systems?
- How is tacit knowledge manifested and exchanged in social-computational systems?

We will implement the research with the Internet Public Library (IPL), a free online digital library with significant STEM content (including K-12 STEM collections). The IPL is accessed by approximately 1 million users per month from around the globe. Providing good quality metadata for the IPL will significantly enhance the usefulness of the IPL for these users, including STEM users. The IPL is an ideal place to test CECM, as it is a distributed organization whose members (IPL students and interns located across the United States) engaged in a complex knowledge-based task (metadata generation).

## TABLE OF CONTENTS

---

For font size and page formatting specifications, see GPG section II.B.2.

	<b>Total No. of Pages</b>	<b>Page No.* (Optional)*</b>
Cover Sheet for Proposal to the National Science Foundation		
Project Summary (not to exceed 1 page)	1	_____
Table of Contents	1	_____
Project Description (Including Results from Prior NSF Support) (not to exceed 15 pages) <b>(Exceed only if allowed by a specific program announcement/solicitation or if approved in advance by the appropriate NSF Assistant Director or designee)</b>	15	_____
References Cited	2	_____
Biographical Sketches (Not to exceed 2 pages each)	13	_____
Budget (Plus up to 3 pages of budget justification)	6	_____
Current and Pending Support	7	_____
Facilities, Equipment and Other Resources	1	_____
Special Information/Supplementary Documentation	0	_____
Appendix (List below. ) <b>(Include only if allowed by a specific program announcement/ solicitation or if approved in advance by the appropriate NSF Assistant Director or designee)</b>	_____	_____
Appendix Items:		

\*Proposers may select any numbering mechanism for the proposal. The entire proposal however, must be paginated. Complete both columns only if the proposal is numbered consecutively.

---

# COLLABORATIVE KNOWLEDGE WORK IN SOCIALLY INTELLIGENT COMPUTATIONAL SYSTEMS

## 1 Introduction

Information systems are a central feature of society. They connect users across space and time, and facilitate work and social collaboration. Information systems have long embodied promises of increased communication, understanding, knowledge and creativity. However their actual use has often played out in emergent, complex and unexpected ways, and central questions, such as what is the nature of knowledge in information systems, and how is knowledge stored and exchanged, remain unresolved. The socially intelligent computational systems of the future will have to address these and other issues in order to succeed. To do so will require sound theoretical and practical understandings of what role knowledge plays in such systems. This proposal pursues such understandings in the context of the emergent behaviors, affordances, unintended outcomes and limits of such systems, through the following questions:

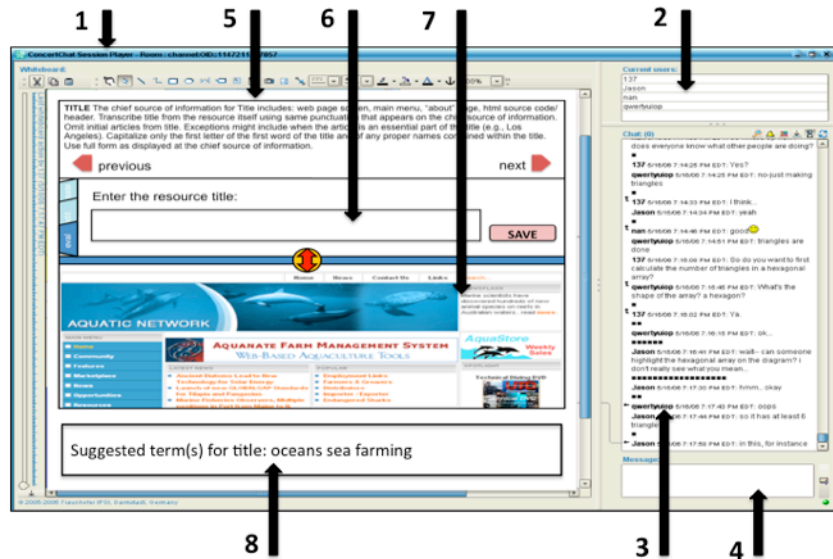
- How can socially intelligent computational systems support users to construct collaborative knowledge?
- How is knowledge shared in socially intelligent computational systems?
- How can complex collaborative work be supported in socially intelligent computational systems and networks?

We investigate these questions through the development and study of a complex collaborative tool to support the creation of high quality metadata for online resources. We directly address a number of the challenges in the SoCS solicitation, including: What methods are effective in studying socially intelligent computing?; How can we better understand what types of behaviors and what new affordances can emerge or be demonstrated by socially intelligent computing?; and, How can we leverage unexpected behaviors of socially intelligent computing systems?

The **intellectual merit** of the proposed work lies in two main areas. First, it addresses a significant unresolved question for collaborative information systems, that is, how can practical knowledge be communicated in such systems, given that that such knowledge itself can be hard to articulate? A second significant subsidiary research question involves gaining understanding of how to support groups of novice metadata creators to acquire the skills necessary to generate low-cost but high-quality metadata for online resources.

The **broader impact** of the proposed work lies in its contributions to understanding how present and future information systems might be designed to share knowledge of complex tasks and support collaborative work. The research also contributes significantly to understanding online resources might be made more accessible to users. This will be done in the context of the Internet Public Library (IPL), a free online digital library with significant STEM content (including K-12 STEM collections). The IPL is accessed by approximately 1 million users per month from around the globe. Providing good quality metadata for the IPL will significantly enhance the usefulness of the IPL for these users, including STEM users. The training of future LIS professionals to use the metadata tool will also contribute significantly to increased metadata capacity in the libraries and digital libraries.

The work will be carried out by an interdisciplinary team who will disseminate the findings widely in relevant research venues, as well as on a project Web site.



**Figure 1a: Conceptual layout of the Collaborative Environment for Creating Metadata (CECM)**

*Existing VMT collaborative environment:*

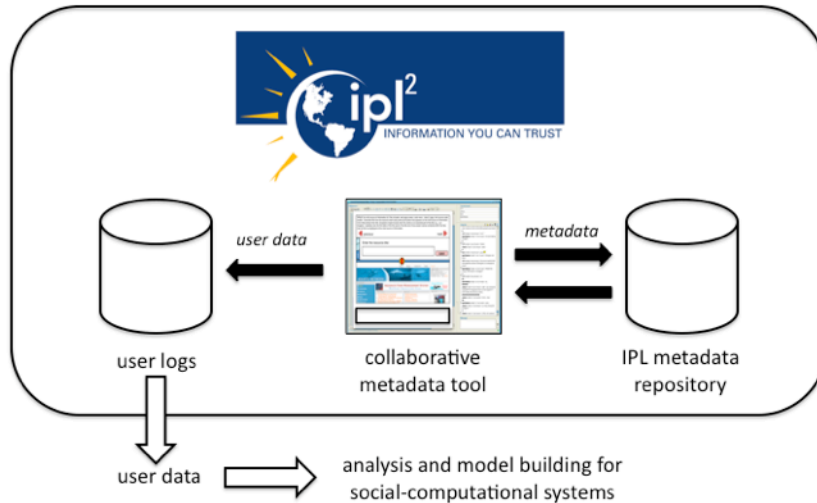
1. VMT window. 2. Users online. 3. Chat window. 4. Chat message composition window

*Metadata administration tool:*

5. Dublin Core rubric and information. 6. Metadata entry window. 7. Web site to be cataloged

*Metadata recommender tool:*

8. Tool will display machine-generated suggestions for metadata terms



**Figure 1b: Conceptual layout of research and data collection**

The IPL will be as a testbed to implement a collaborative metadata generation tool (see Figure 1a, above). The tool will generate metadata for the IPL, and user logs and data for the analysis of the wider research questions regarding knowledge in socially intelligent computational systems.

## 2 Significance and Need

Computer-based information systems connect users instantaneously across time and space. They promise to improved communication, work and social collaboration, and creativity and productivity. However, this promise can play out in emergent, complex, and unexpected ways, at societal, organizational, and individual levels. Examples of the complexity of information system adoption include the 'productivity paradox' of 1980s office automation (Brynjolfsson 1993; Brynjolfsson & Hitt, 1998); debates surrounding the efficacy (or otherwise) of educational technologies (e.g. Khoo 2007); and emergent links between mobile device use and social activities such as driving (e.g. Strayer et al. 2006).

These and many other examples show us that models and explanations of information system use have to be considered within the context of human behavior. From a science, technology and society (STS) perspective, socially intelligent computational systems, like all technologies, are complex systems of people, technologies, and economic, historical, and other components, and the iterative interactions between these multiple components result in unexpected social and technological outcomes. Current and proposed plans for cyberinfrastructure and socially intelligent computational systems make the STS approach more relevant than ever. These new systems will have increased connectivity, power and bandwidth, will allow users to create and share content in new and innovative ways, and these will also increase the complexity of the attendant technological and social processes. The analysis and modeling of these systems and components will require a deep understanding of how human intelligence and knowledge are mediated in and amongst groups in complex distributed computational environments.

**We propose therefore to study the relationships between human, knowledge, organizational, and computational elements in a prototypical social-computational system (the Collaborative Environment for Creating Metadata: 'CECM'), with the aim of understanding how such a system supports communities of practice, tacit knowledge, and legitimate peripheral participation. The study will provide support for further inquiries into the requirements, affordances, and capabilities of such systems.**

Through a detailed investigation of one specific case the study will provide support for further inquiries into the knowledge requirements, affordances, and capabilities of such systems. The study will integrate both social and computational components. The social component includes models of the behavior of individuals and groups when engaged in online collaborative work, and the social dimensions of knowledge sharing in online settings. The computational component includes the development of social network workspaces and automatic recommender tools to support knowledge work and metadata generation by non-experts. The integration is built on analyses of communication and knowledge construction in these tools at both individual and group levels, using a range of different analytical techniques, including text and data mining, content analysis, discourse analysis, user metrics and others. It will study the interactions between the social and technical components, and will address central social-computational research questions such as:

- **How can socially intelligent computational systems support users to construct collaborative knowledge?**
- **How is knowledge shared in socially intelligent computational systems?**
- **How can complex collaborative work be supported in socially intelligent computational systems and networks?**

Through a detailed investigation of one specific case – the use of CECM – the study will provide support for further inquiries into the knowledge requirements, affordances, and capabilities of such systems.

### 3 Theoretical Approach

CECM (described below) will provide a test bed for an analysis of knowledge in a prototypical social-computational system. Our analysis is based on three important social models of collaborative work. These models are:

- *Communities of practice*, groups of people engaged in a common task and who share knowledge of that task amongst themselves
- *Tacit knowledge*, practice-based knowledge that cannot be articulated, but which plays a crucial role in community knowledge and creativity
- *Legitimate peripheral participation*, the process by which community members learn community knowledge

Wenger's (1998) theory of *communities of practice* describes how groups of people engage in a common task and share knowledge of that task amongst themselves. A community of practice is a "collection of individuals sharing mutually defined practices, beliefs, and understandings over an extended time frame in the pursuit of a shared enterprise." In the context of this proposal, this could include knowledge of how to generate metadata for collections of online resources. However, while such mutual understandings support the functioning of the community they are not necessarily obvious to outsiders, and have to be learned over time. They include tacit knowledge, knowledge that we use to support everyday practice but which is hard to articulate; such knowledge often appears new and complex to those joining a community of practice, while at the same time, existing members of that community might have trouble articulating that knowledge in a concise way to new members.

An important part of the knowledge in communities of practice is *tacit knowledge*, knowledge that is practice-based and is hard to articulate (Polanyi 1967). It is often contrasted with explicit knowledge, which can be articulated and written down (Choo 1998; Nonaka & Takeuchi 1995). Tacit knowledge includes 'know-how,' the practical expertise that allows a task to be accomplished, such as riding a bicycle, or playing a musical instrument, rather than 'know-what,' such as information about how a bicycle works. In the context of this proposal, this could include knowledge of how to decide *how* to catalog a resource with a metadata tool, rather than what metadata are. For instance, when cataloging a Web site, we know that the title of a resource is the piece of information that describes the site overall, but is the title of that site the information in the HTML <meta> tag, or the text of the largest/first heading (as these often differ), or some other information on the Web page? Tacit knowledge in this case involves making a complex situated judgment regarding the most appropriate choice.

Tacit knowledge is learned over time in communities of practice when novices learn from more experienced members, practice their new knowledge, and become in time more experienced members themselves who are able to support the induction of new novices. Over time, a group's tacit knowledge can become taken-for-granted by its members, and hard to articulate. Lave and Wenger's (1991) model of *legitimate peripheral participation* describes how learning within communities of practice is facilitated when novices engage in conversation with more experienced members, becoming in time more experienced members themselves who are able to support the induction of new novices. In the context of this proposal, this could include people with more experience of metadata generation assisting those with less or no experience. The

process depends on new members of a community of practice having access to experts who can help and advise them.

Taken together, these theoretical frameworks help us to understand how people in groups, define their tasks, and share knowledge about and accomplish those tasks, as well as how knowledge in groups is accumulated over time and transferred from experienced to novice members. They generate three further specific research questions for the proposed work:

- How do communities of practice form in socially intelligent computational systems?
- How does legitimate peripheral participation occur in socially intelligent computational systems?
- How is tacit knowledge manifested and shared in socially intelligent computational systems?

## **4 The Collaborative Environment For Creating Metadata (CECM)**

We will implement our theoretical approach in the context of a specific technical test-bed: groups of graduate students engaged in a online task, the creation of metadata for digital resources, using a tool that we will develop, the Collaborative Environment for Creating Metadata (CECM). CECM represents a prototypical socially intelligent computational system because it combines human and machine intelligence in a sophisticated collaborative environment, with the aim of generating complex knowledge artifacts.

Note that while building the CECM test-bed is an interesting research activity in itself – it addresses a central concern of cyberinfrastructure, that of how accurately to locate good quality information resources on the Internet – it is not the central subject of this proposal. Rather, it will serve as an environment that will support the gathering of data to answer our wider research questions regarding knowledge in socially intelligent computational systems. In order to achieve our research goals in a timely fashion, we will therefore adapt and develop existing tools rather than build our own test-bed completely from scratch. CECM will therefore be developed from three existing components: the Virtual Math Teams (VMT) environment of the Math Forum, a cataloging tool currently under development at the Internet Public Library (IPL), and a prototype automatic metadata extraction tool, also currently in development at the IPL.

### **4.1 Background: The Metadata Bottleneck**

CECM addresses a long-standing issue in the area of online resources, that of how to search for and retrieve high quality information from the Internet. Computer networks and information and communication technologies have facilitated rapid access to large amounts of information. At the same time they have also supported the creation of further vast amounts of information that in turn has to be organized, indexed, and made searchable. Here, one approach involves the creation of metadata. Metadata are abstracted descriptions of an information resource – for instance in the form of a catalog record – that can be searched in order to locate a particular resource. For example, the catalog record for a Web site that describes the life of George Washington could include such terms as ‘George Washington,’ ‘president,’ ‘United States of America,’ and so on, as well as the URL of the site, a short description or abstract, and other information. Good quality metadata support rich user interactions with a repository, while poor quality metadata (e.g. ‘George Wasnigton’) hide resources, produce poor search results, and negatively affect user satisfaction (Barton et al. 2004; Beall 2005; Geisler et al. 2002).

There is currently a ‘metadata generation bottleneck’ between the growing numbers of digital resources requiring description, and the limited numbers of metadata experts who create such

descriptions (Liddy et al. 2001). Approaches to addressing this bottleneck include training metadata specialists, a long-term and expensive solution; automatic metadata extraction (e.g. Greenberg et al., 2005), an approach which can have varying degrees of success, depending on the quality of the resources being indexed; and supporting and training non-specialists to create metadata. CECM will focus on this last strategy.

The use of non-specialists to generate metadata is an attractive proposition in terms of resource allocation, as it avoids expensive cataloging. However, it has had mixed success in terms of productivity and of the quality of the metadata generated. While the task of describing a resource such as Web site within a narrow set of guidelines and/or controlled vocabularies might appear to be a relatively easy one, it has proved to be difficult for many users, even when supported by sophisticated tools and training (Khoo 2005). The reported problems have been attributed variously to the complexity of the metadata concepts that the tools convey, poor guidelines, poor interface design, a lack of project resources, etc. (Crystal and Greenberg 2005; Greenberg et al. 2003; Kastens et al. 2005; Lagoze et al. 2005; Wilson 2007). A problem of relevance to this proposal has been the failure to communicate knowledge about metadata creation to and amongst the users of an online metadata tool (Khoo 2005).

## **4.2 The Collaborative Environment For Creating Metadata (CECM)**

As part of our investigation into the sociotechnical dimensions of socially intelligent computational systems, we will assume that the quality of metadata can be improved by supporting metadata creators in online communities of practice, embedded in a collaborative interaction space – the Collaborative Environment for Creating Metadata (CECM).

CECM will support groups to generate metadata records in real time, to observe each other doing so, and to engage in real-time discussion about their task. It will allow new or novice users to learn about metadata from more experienced users, and it will support more experienced users to instruct and guide new users in these areas. Besides metadata creation and editing, CECM will support simultaneous real-time viewing of cataloging tool use by multiple users, chat channels for users to discuss their common tasks in real time, and wikis for meta-communication.

CECM will consist of three major components: an online collaborative environment, and two metadata tools, one to support metadata record editing and creation, and one to provide automatic metadata extraction recommendations (see figure 1a). These components will be integrated in novel ways that will allow users to create metadata, access machine-generated metadata recommendations, share metadata knowledge, and watch each other doing so, all in a linked textual-graphical environment. CECM is a sophisticated tool that attempts to leverage machine-based and human information and knowledge capabilities in pursuit of a real-world goal (cataloging).

For the purpose of this SoCS proposal, we argue that CECM represents a prototypical social-computational system that is realizable now with existing tools and tool components; and that it will provide a useful test-bed for investigating some of the sociotechnical and knowledge-based dimensions of socially intelligent computational systems.

### *4.2.1 The Virtual Math Team (VMT) Environment of the Math Forum*

The collaborative environment for CECM will be based on the existing Virtual Math Team (VMT) tool that is part of the Math Forum (<http://www.mathforum.org>), a Web site with over a million pages of resources related to mathematics for middle-school and high-school students, primarily



algebra and geometry. The Math Forum is visited by several million different visitors a month, including teachers, mathematicians, researchers, students and parents.

The VMT service currently consists of an introductory web portal within the Math Forum site and an interactive software environment. The VMT environment includes the VMT Lobby, where people can select chat rooms to enter, and a variety of math discussion chat rooms, that each include a text chat window (on right), a shared drawing area and a number of related tools (on left) (see figure 2) (Stahl 2009).

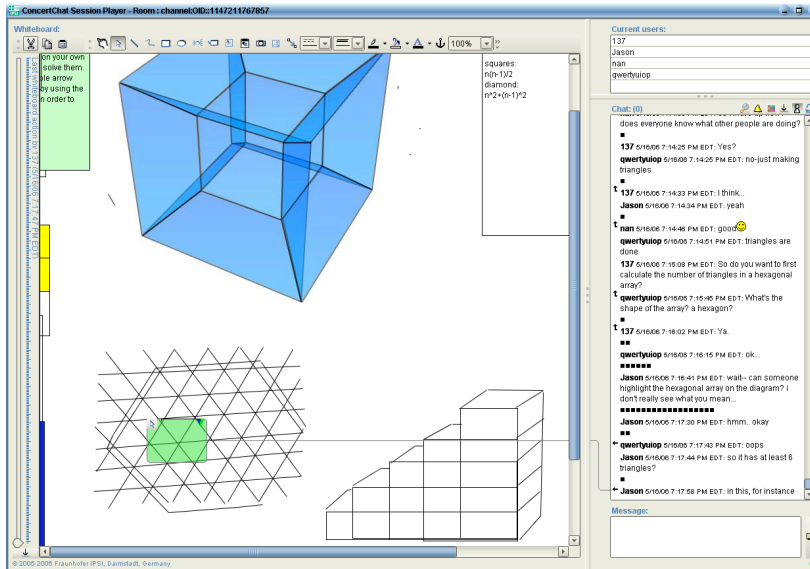


Figure 2: The VMT tool

A key feature of the tool is the real-time interaction between the shared drawing area and the chat environment. In this environment, students can chat online about the mathematical problem that they are addressing, while at the same time communicating graphically using the shared whiteboard space. This affordance of VMT allows students to 'show' each other what they are talking about, and helps them to discuss, learn, and absorb complex mathematical concepts and problems in a collaborative fashion.

VMT was designed to support students in the collaborative learning of mathematics. We will adapt the tool to the collaborative generation of metadata in CECM by integrating two new components into VMT. These components will take the place of the whiteboard tool on the left side of the current VMT tool (compare figures 2, and 1a). One component will be the next version of a cataloging tool currently under development with the IPL, and the second component will be an automatic metadata recommender tool. We will keep the current chat setup in VMT. Together, these changes mean that groups of users will be able to use the cataloging tool and metadata recommender, and discuss and illustrate their use of these tools in the chat channel.

#### 4.2.2 Automatic Metadata Suggestion Tool

Besides supporting real-time collaboration of catalogers in metadata generation, a further innovative feature of our tool is to provide suggestions to support metadata creation. Suggestions of metadata can be automatically generated by extracting semantic relations among multiple related sources created by users either individually or collaboratively. Different sources capture the semantic context of a collection from a different perspective. Some potential semantic sources include keywords manually assigned by catalogers, tags manually assigned by regular users, query keywords identified and selected from previous user search logs, and so

on (Lin et al., 2008). We particularly explore how context might be employed for metadata and how the context information might be extracted from both the semantic analysis of digital collections and the analysis of user's search logs.

In order to extract semantic relations from multiple sources, topic signatures are a key concept. The topic signatures can be automatically generated through a topic signature model we developed (Zhou et al. 2006). The model is based on semantic mapping through a language modeling approach and a context-sensitive semantic smoothing method (Zhou et al., 2007). Two types of mappings are created in this language model. One is called *topic signatures* that map from any term,  $w$ , in the collection to a list of topics,  $t$ 's (represented by keywords, subject headings, or other indexing terms). The other is called *semantic profiles* that map a specific topic ( $t$ ) to a set of terms ( $w$ 's) that are most likely to co-occur with  $t$  in the collection.

Given a collection ( $C$ ), we first index all documents with individual terms and topics. For each topic  $t_k$ , we approximate its semantic profile using the terms  $w$ 's in the document set  $D_k$  containing  $t_k$ , ranked in the descending order of the conditional probability  $p(w | t_k, C)$ . We assume that the terms appearing in  $D_k$  are generated by a mixture model:

$$p(w | t_k, C) = (1 - \alpha) p(w | t_k) + \alpha p(w | C)$$

where  $p(w | t_k)$  is a topic model that represents the conditional probability of term  $w$  co-occurring with topic  $t_k$ .  $p(w | C)$  is a background model describing the global distribution of terms in the collection  $C$ , and  $\alpha$  accounts for the background noise. Not only does this mixture model capture the semantic associations between topics and terms in the topic model, but it takes into account the overall term distribution of a collection in the background model. The model for  $t_k$  can be estimated using an expectation-maximization (EM) algorithm. Details of the model can be seen in (Zhou et al., 2006; Zhou et al., 2007). It is worth noting that the model represents an effective semantic mapping based on both the collection content and its context. The metadata used to describe the same terms or subjects may vary from collection to collection. Our model is able to capture the different semantic associations among topic signatures in different collections.

With the language model in the back-end, our tool will have an interface that functions as a "semantic aggregator" and a collaborative authoring workspace that provides access to multiple semantic sources, including the metadata, topic signatures, semantic profiles, and most frequently used search terms. The interface provides rich interactive functions and links. It allows user to create multiple types of metadata such as subject terms, topics, themes, and so on. Keywords or terms automatically extracted by the semantic mapping model will be suggested to users for metadata creation. Users can easily select or edit generated terms to enhance the representations.

#### 4.2.3 User-friendly metadata interface

The third component of the tool is a metadata administration tool that permits the viewing, creation and editing of metadata records. An initial version of this tool is currently under development, funded by a grant from OCLC, and shaped by several important design criteria. First, as the administration of IPL metadata records currently requires multiple browser windows to accomplish this task, including one for the record, one for the online resource, one for the evaluation rubric, and one for the evaluation form, the new interface will integrate these functions within a single window (c.f. Khoo et al. 2002). Second, the interface development is following an iterative user-centered HCI design process to avoid usability problems (e.g. Nielsen 2005; Norman 2002). Third, even a metadata tool that is carefully designed from an HCI perspective can be problematic for non-experts to use, because of the ways in which that tool

represents technically complex metadata concepts to the tool users (Khoo 2005). Therefore, close attention is being paid to the design and placement of the metadata evaluation rubrics and forms, and associated 'help' and FAQ information, etc., that will support IPL students and volunteers to carry out their tasks.

For this proposal, we will integrate this tool into the VMT environment and refine it for use in collaborative group settings.

#### *4.2.4 Integration*

The metadata cataloging and recommender tools will be implemented in a modified VMT environment. There are interesting (but hopefully not too substantial) issues involved with the integration of these tools, including making sure that the VMT environment can communicate with the current Fedora backend of the IPL (Fedora is an open source repository tool: <http://www.fedora-commons.org/>). We will employ a programmer to work on these integration issues.

### **4.3 Implementation**

We will implement CECM within the Internet Public Library (IPL: <http://www.ipl.org>). The IPL was created in 1995 at the University of Michigan-Ann Arbor, and it is now available throughout the United States as a training tool for library and information science (LIS) programs. It has trained over 5,000 students in 17 LIS programs in tasks such as the creation and editing of metadata. The IPL has subject-categorized collections of more than 40,000 online resources. These collections are the cumulative result of work done by various students, volunteers, and staff members. A recent pilot study of the quality of the metadata in the IPL (as well as in a partner library, the Librarian's Internet Index, LII: <http://www.lii.org/>), showed that the catalog records of both digital libraries contained metadata that varied widely in quality (see Wilson 2007 for descriptions of metadata quality measures). The expectation prior to carrying out this study was that metadata would be relatively consistent, although given that many records were created by trainees, not completely so. The inconsistencies found in both libraries' metadata were attributable in part to the 'piecemeal' development of both libraries by a large number of distributed individuals.

The IPL is therefore an ideal place to test CECM. It also represents a distributed organization (IPL students and interns are located across the United States), whose members are engaged in a complex knowledge-based task (metadata generation), with participants of varying degrees of expertise. If CECM represents a prototypical social-computational system, then the IPL can be said to represent one kind of a prototypical social-computational task.

CECM will be used to support metadata generation assignments in graduate library and information science classes Drexel University and other universities of the IPL consortium. These classes are expected to provide 50-100 users per year. As part of this implementation, we will develop appropriate course modules to support instructors to implement CECM in their classrooms, and also to support students to use the tool. We will also involve instructor and student users in the iterative design and improvement of CECM by inviting them to submit feedback on the tool's design and functionality. (For further details, see work plan, below).

## **5 Data Collection and Analysis**

The CECM test bed will be used as a platform for gathering data to address the central research questions of this proposal regarding the creation and sharing of knowledge in a prototypical social computational system. A crucial aspect CECM will be the facility for integrated data

collection across all dimensions of the tool. We will be able to record how users interact with the tool and with each other, and follow their chat conversations as they engage in use of the metadata administration and the metadata suggestion tool, and view the metadata that they create. In other words, *we will be able to track what was said, when it was said, who it was said to, what was being done with the metadata tools and what metadata were being created at the time.*

All data streams will be time-stamped, and we will be able to correlate individual and group actions, group/social network membership, individual and group discourse, knowledge work (metadata records), etc. The data will be used to address our specific research questions (How do communities of practice form in socially intelligent computational systems? How does legitimate peripheral participation occur in socially intelligent computational systems? and, How is tacit knowledge manifested and shared in socially intelligent computational systems?). The data analysis will include:

- Patterns of group and social interaction (content analysis and network analysis)
- Patterns of individual and group knowledge (discourse coding and computational text analysis)
- Patterns of tool use (user metrics)
- Metadata quality

These approaches will be applied to the data collected from CECM, and the results will be triangulated in order to identify and confirm the key concepts present in these data.

## **5.1 Content Analysis and Network Analysis**

We will analyze the interaction data generated by the groups collaborating in CECM using content analysis (Krippendorf, 2004) and network analysis (Wasserman and Faust, 1994). The content analysis will be performed on the all text chat postings and the network analysis will be performed on the same chat postings.

The content analysis will be executed using two rubrics. The unit of analysis for this work will be a complete unit of conversation (Krippendorf, 2004). The first rubric will evaluate the development of group identity within the small groups, using Tajfel's (1978, 1979, 1982) description of group communication as inter-group, inter-personal, intra-group and inter-individual. Inter-group communication is communication across groups, and only rarely occurs in this data set. Inter-personal communication takes place between two individuals. Intra-group communication is within the group, where all members participate in the dialogue. The addressing individual members in the presence of the whole group as an aside is coded as inter-individual communication. The second rubric will evaluate the corpus of data for knowledge co-construction using a rubric developed by Gunawardena et al (1997). Two raters will score the conversations on these rubrics and measure inter-rater reliability using Krippendorf's alpha (2004). This type of analysis is performed on asynchronous communication records, and the contrast with the results from synchronous chat data will provide a helpful contrast of synchronous and asynchronous knowledge co-construction in small groups.

Social network analysis will be performed on the chat postings and other actions in order to determine if there are patterns of networked interaction that correspond with the development of group identity or the co-construction of knowledge. The resulting networks will be bi-partite (users and objects) and regular. Since the networks in our corpora are closed and small, we will focus our analysis on small network evolution and elaborating semantically meaningful measures of tie strength. Tracking longitudinal evolution will involve developing a time-series set

of network views, possibly addressing the state of the network as a feature that contributes to the other forms of analysis. We will also explore the advantages of deriving measures of tie strength from the results of machine learning algorithms, response time lag and length of sustained interaction between pairs of group members.

These quantitative analyses will not be performed in isolation from the interaction analysis or the automated coding. Decisions about the granularity in both network analysis and content analysis will take the findings and approaches from these other two methods into consideration. The findings of all these mixed-method analyses will inform the design of computational models and supply a basis for calibrating the models of macrocognition.

## **5.2 Discourse Coding and Computational Text Analysis**

Our second analytical approach will look for emergent patterns of individual and group knowledge. This will be done primarily through the analysis and coding of the discourse of the users of CECM. For instance, we might look for evidence of acquired understanding and knowledge of metadata, by comparing users' chat in the tool at the start of a metadata class with the chat at the end of the class. Changes in vocabulary of users (for instance, in the terms that they use to describe metadata, or in the practices and actions that they associate with metadata work) could indicate evidence of having acquired understanding of metadata. Conversely, lack of changes in vocabulary could be markers that those understandings were not acquired (Khoo 2005).

One approach will therefore be to categorize and code the chat data through an inductive, grounded theoretical approach (Glaser & Strauss 1967). In this approach the collected documents will be reviewed and marked up for significant concepts. The initially identified concepts will be recorded, and used to guide subsequent iterative rounds of coding of the same documents. During these latter stages the initial categories will be adapted and/or refined, and new categories may be added, before a final set of concepts is identified. Off-the-shelf software exists (e.g. N-VIVO: <http://www.qsrinternational.com/>) that can be used to support this coding.

The same chat documents will also be analyzed with a computational text analysis tool, Centering Resonance Analysis (CRA) ([www.crawdadtch.com](http://www.crawdadtch.com)). CRA assumes that "communicators speak or write coherently by creating utterances that deploy a stream of words comprising centers (more specifically, noun phrases) in a strategic way, creating a semantic structure of words" (Corman et al. 2002). CRA parses documents into noun phrases, and then calculates an index of significance for each noun based on the frequency of its occurrence, and also its co-occurrence in relation to all the other nouns in the text. These results can be viewed as a ranked list of significant single and co-occurring nouns, and also as an image file (essentially a graph of nodes and edges) of the nouns and the relationships between the nouns. The tool can also cluster documents based on similarities in their individual CRA analyses.

While discourse coding and computational text analysis represent different approaches, they are not mutually exclusive, and can provide strong evidence for the presence or absence of particular forms of knowledge, especially if they are triangulated.

## **5.3 User Metrics and Patterns of Tool Use and Task Work**

The CEMC will be fully instrumented to supply data of all user interactions with the environment, including chat postings and use of the metadata tool. We will therefore be able to triangulate users' actions, and the content, social network, and discourse analyses, with their actual actions in the environment, such as the use of the metadata tools and the creation of metadata. Here again this can provide for strong evidence of metadata knowledge, for example, if we can

correlate a particular section of chat with a particular set of actions in CECM. For instance, one user could explain a metadata issue to a second user; does the second user then start to produce better quality metadata?

#### **5.4 Metadata Quality**

Our final dataset will consist of the metadata generated by the users over the course of the project. We will subject these metadata records to two forms of analysis: (1) a quantitative analysis of the completion of various metadata fields, and (2) a qualitative analysis of the content of these fields, based on a random sample of records from the repository, submitted for analysis to metadata experts. We will be able both to assess the overall quality of metadata produced within CECM, and also attempt to identify any longitudinal trends (for instance, do metadata get better in quality over time as users talk about the issue involved?).

#### **5.5 Data Triangulation**

We will collect quantitative and qualitative data across both the social and machine dimensions of CECM. We will analyze these data using a variety of quantitative and qualitative techniques. We will then use the results of these analyses to begin the development of theories and models of knowledge work in socially intelligent computational systems, based on existing theories of communities of practice, tacit knowledge, and legitimate peripheral participation. Triangulation of methods is a crucial feature of the data analysis that can address this complexity. For example, one important route of inquiry will be to compare (a) the quality of metadata records, (b) the collaborative context in which these records were created – were they created by individuals or groups? (social network analysis), and (c) the vocabulary used by the participants – are expert metadata terms used? (discourse analysis, computational text analysis). This particular route of inquiry will help us to understand if and how high quality metadata records are created, whether or not this is done in the context of groups, whether or not these groups constitute communities of practice, and whether or not they generate and share expert knowledge and terminologies.

### **6 Work Plan**

#### **Year 1 - Development**

We will begin the adaptation and development of the VMT environment discussed in section 3. This will include the development of the metadata suggestion and administration tools and their integration into VMT. We will also integrate the new CECM front-end with the existing back-end of IPL (currently a Fedora database: <http://www.fedoracommons.org/>). A prototype tool will be available for early user testing in the second half of 2010. It will be tested in HCI classes taught by Khoo, where it will serve as a real-life 'case study' for HCI students, as well as in metadata assignments taught by Abels and Lin at Drexel. We will also begin the development of the learning modules that will support CECM use in the classroom, including pedagogical materials, tool instructions, assignments, and assessment sheets. The modules will also be added to the Virtual Learning Laboratory for Digital Reference (VLL) at the IPL (VLL: <http://vll.ipl.org>), a collection of learning objects for digital reference and collection development that are available to instructors in LIS programs for use in their classes.

#### **Year 2 – Implementation**

The prototype version of CECM will be implemented in cataloging classes at Drexel and at IPL consortium universities (see: <http://ipl.org/div/about/IPLconsortium/consortiumList.html>). Evaluation of the tool will begin (see below). As CECM is implemented, we will start to collect and analyze the data necessary for addressing this proposal's overall research questions, and

looking for emergent and unexpected patterns of human and machine behavior. These data will include chat logs, metadata tool use, metadata records created, etc.

### Year 3 – Analysis, Evaluation, Dissemination

We will continue data collection and analysis, and start work towards generating models of the knowledge activities that have been recorded. We will do this in regular project meetings and data sessions. We will also begin summative evaluation activities (see below). Finally, we will prepare and submit articles documenting the project and outcomes to journals and conferences within the PI’s and Co-PI’s individual areas of expertise.

		2010 - Development				2011 - Implementation				2012 – Analysis & Eval.			
		Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
<b>CECM</b>		Initial development				Ongoing development and refinement							
			Early user testing										
			Class implementation										
			Class-based user evaluation										
			Documentation and dissemination										
			Summative evaluation										
			Instructional module development										
<b>Project</b>		Data sessions and analysis											
		Theory and model building											
		Writing											
		Dissemination											

Figure 3: Overall timeline

## 7 Outcomes and Evaluation

There are two high-level sets of outcomes for this project:

- Primary: Analyses of CECM user data leading to the development of a model of knowledge work in socially intelligent computational systems
- Secondary: Development and implementation of CECM, metadata generation

These two sets of outcomes correspond approximately to formative and summative levels of evaluation. Formative evaluation activities will take place at all stages of the project, to assess whether the goals of individual stages have been met, and to what extent the processes in each stage are proceeding as planned.

The primary outcome will be assessed through summative evaluation. This will take place at the end of the project, and will assess the extent to which the project has met its stated goals. In the case of this proposal, the goals are for theory generation for socially intelligent computational systems. We will therefore write, from the perspective of the various computational, social scientific, and other expertises of the project team, a series of research papers detailing our findings and theoretical insights, which will be submitted to academic conferences and journals conferences. In this case, the evaluation will be provided by the peer review processes of these venues.

Second, there numerous subsidiary outcomes of the work, many related to CECM and its use. Although it is a central theoretical assumption of this proposal that many of these outcomes will be unexpected, some of the outcomes that are anticipated include:

- Development of a usable tool
- Development of curricula and learning objects to support the use of CECM
- Student understanding of metadata
- Student and instructor satisfaction with instructional units
- High quality metadata

These will be assessed as the project progresses. CECM usability will be assessed through testing in HCI classes run by Khoo at Drexel, and also with students in metadata classes. The usefulness and usability of CECM for professors and students will be assessed through in-class surveys and in-class discussions on electronic discussion boards. Feedback will be collected on the tool's features, the efficacy of the tool as an enhancement to learning, etc. Student understanding of metadata concepts will be derived from the analysis of the chat and other VMY logs, and also through in-class surveys at the beginning and end of courses. Finally, the quality of the metadata generated in CECM will be analyzed through quantitative analysis of metadata fields, and a qualitative analysis a random sample of records from the repository by metadata experts (see 5.4 above).

## 8 Dissemination

The PI and co-PIs have extensive ties with multiple relevant fields of research, and will collaboratively author publications, presentations, and demos to be submitted to relevant journals and conferences in their relevant fields of specialization (see the bios of each project member for further details). The grant includes a total of 3 domestic trips and 1 international trip each for Khoo, Goggins and Li, to support these dissemination activities.

In addition, we will create a project Web site to describe our work that will contain project information, research bibliographies, and reports of ongoing findings and research results.

## 9 Personnel and Results from Prior Support

### 9.1 Personnel

The investigators bring a highly relevant and synergetic blend of interdisciplinary research skills and interests to this proposal. Khoo, Goggins, and Li will carry out the majority of the work. **Khoo** has nine years' experience working with digital libraries, metadata tools, and novices creating metadata, and his research interests include communities of practice and tacit knowledge. He is currently recipient of an OCLC-ALISE grant to develop a metadata administration tool for the IPL. **Goggins** focuses on the development of socio-technical systems that support small group collaboration. His expertise includes mixed methods research focused on social interaction in online groups, the explication of completely online groups as an emerging phenomena in socio-technical research, and the use of network analysis to extend and develop theories of group cognition. **Li** has extensive experience in data mining and text mining techniques. He has worked on a prototype of a semi-automatic metadata tool for the IPL. His expertise in text analysis and online communication analysis is of critical importance to the success of the proposed research.

With extensive experience in 3 critical areas of the project, Abels, Lin, and Stahl will offer additional support over the course of the proposed work, including liason between the IPL and the Math Forum/VMT. **Eileen Abels** has been overseeing the operations of the Internet Public Library since it was moved to Drexel in January 2007. Prior to that, she taught in the area of information access and digital reference. Her current research interests provide a bridge between practice and LIS education. **Xia Lin's** research areas include digital libraries, semantic content analysis, information visualization, and visual interface design. He is particularly familiar



with the IPL system and content architecture. **Gerry Stahl** conducts research in human-computer interaction (HCI) and computer-supported collaborative learning (CSCL). He is the Principal Investigator of the Virtual Math Teams Project, and will provide a crucial bridge between the MathForum developers and the IPL.

Together the PI, co-PIs and Senior Personnel are able to address both the computational and the social dimensions of the proposed research in a synergetic and interdisciplinary fashion.

## **9.2 Results From Prior Support**

### Xia Lin

*Planning Grant: I/UCRC Center for Visual Decision Informatics* (award #: 0934197, amount \$10,000, 09/01/09 – 08/31/10). The grant supports planning and operational activities that will lead to the establishment of a consortium of universities and industry partnerships on visual information processing and decisions making.

### Gerry Stahl

*IERI: Catalyzing & Nurturing Online Workgroups to Power Virtual Learning Communities* (award # 0325447, \$2,300,000, 09/01/03 - 08/31/09. PI: Gerry Stahl; co-PIs: Stephen Weimar and Wesley Shumar. The grant successfully completed several iterations of design, development, testing and analysis of the VMT online collaborative math service at the Math Forum. Over 1,000 student-sessions took place, averaging an hour each. Over 150 publications associated with this project have appeared. Software for the VMT environment is being released as Open Source and is being used by other researchers in collaboration with on-going work continuing from this project (Stahl, 2006; 2009).

## **10 Summary**

The proposed research will generate a better understanding of the social and computational dimensions of a complex collaborative task (metadata generation) in a prototypical social-computational system. The proposed study is a case study of emergent relationships between human, organizational, and computational elements, which will provide the building blocks for the development of wider theoretical inquiries aimed at understanding the requirements, affordances, and capabilities of socially intelligent computational systems.

## References

- Barton, J., Currier, S., & Hey, J. (2003). Building Quality Assurance into Metadata Creation. Dublin Core Conference 2003, 28th Sept - 2nd Oct, Seattle, Washington, USA.
- Beall, J. (2004). Metadata and data quality problems in the digital library. *Journal of Digital Information* 6(3).
- Brin, S., and Page, L. (1998). The anatomy of a large-scale hypertextual search engine. *Computer Networks and ISDN Systems*, 30 (1998), pp. 107-117.
- Brynjolfsson, E. (1993). "The productivity paradox of information technology", *Communications of the ACM*, 36(12), pp. 67-77.
- Brynjolfsson, E., and Hitt, L. (1998). "Beyond the productivity paradox. Computers are the catalyst for bigger changes," *Communications of the ACM*, 41(8), pp. 49-55.
- Choo, C. (1998). *The Knowing Organization*. Oxford: Oxford University Press.
- Crystal, A., & Greenberg, J. (2005). Usability of a metadata creation application for resource authors. *Library & Information Science Research* 27 (2005), 177-189.
- Geisler, G., Giersch, S., McArthur, D., & McClelland, M. (2002). Creating virtual collections in Digital Libraries: Benefits and implementation issues. JCDL '02, Portland OR, USA, July 13-17, 2002. ACM, New York, NY, pp. 210-218.
- Glaser, B. & Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago, IL: Aldine Publishing Company.
- Greenberg, J., Crystal, A., Robertson, W., & Leadem, E. (2003) Iterative design of metadata creation tools for resource authors. 2003 Dublin Core Conference 'Supporting Communities of Discourse and Practice,' Seattle, WA, USA.
- Greenberg, J., Spurgin, K., & Crystal, A. (2005). Final Report for the AMeGA (Automatic Metadata Generation Applications) Project.  
[www.loc.gov/catdir/bibcontrol/lc\\_amega\\_final\\_report.pdf](http://www.loc.gov/catdir/bibcontrol/lc_amega_final_report.pdf)
- Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of Global Online Debate and the Development of an Interaction Analysis Model for Examining Social Construction of Knowledge in Computer Conferencing. *Journal of Educational Computing Research*, 17(4).
- Kastens, K.A., DeFelice, B., Devaul, H., DiLeonardo, C., Ginger, K., Larsen, S., Mogk, D., & Tahirkheli, A. (2005). Questions & challenges arising in building the collection of a digital library for education: Lessons from five years of DLESE. *D-Lib Magazine* 11(11).  
<http://www.dlib.org/dlib/november05/kastens/11kastens.html>.
- Khoo, M. (2005). The Tacit Dimensions of User Behavior: The Case of the Digital Water Education Library. 5th ACM/IEEE Joint Conference on Digital Libraries (JCDL), Denver, CO, June 7-11, 2005, p. 213-222.
- Khoo, M. (2007). Observing the users of digital educational technologies – theories, methods and analytical approaches. Editorial. *New Review of Hypermedia and Multimedia*, 13(2), December 2007.
- Khoo, M., Devaul, H., and Sumner, T. (2002). Functional Requirements for Groupware to Support Community-Led Collections Building. 6th European Conference on Digital Libraries, Rome, September 2002, pp. 190-203.
- Khoo, M., Lin, X., & Park, J (2009). *A User-Friendly Metadata Quality Control Tool for the Internet Public Library*. 9th ACM/IEEE Joint Conference on Digital Libraries (JCDL), Austin, TX, June 15-19, 2009, pp. 407-408.
- Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology*. Thousand Oaks, CA: Sage Publications.

- Lagoze, C., D. Krafft, T. Cornwell, N. Dushay, D. Eckstrom, & J. Saylor. (2006). Metadata aggregation and "automated digital libraries": A retrospective on the NSDL experience. JCDL '06, pp. 230-239. ACM Press.
- Lave, J., & Wenger, E. (1991). *Situated learning*. Cambridge: Cambridge University Press.
- Liddy, E., Sutton, S., Paik, W., Allen, E., Harwell, S., Monsour, M., Turner, A., & Liddy, J. (2001). Breaking the metadata generation bottleneck: preliminary findings. JCDL '01, p. 464. ACM Press.
- Lin, X., Li, J., & Zhou, X. (2008). Theme creation for digital collections. Proceedings of the International Dublin Core and Metadata Applications (DC 2008), 34-42, September 22-26, Berlin, German.
- Nielsen, J. (2005). Ten usability heuristics.  
[http://www.useit.com/papers/heuristic/heuristic\\_list.html](http://www.useit.com/papers/heuristic/heuristic_list.html)
- Nonaka, I., & Takeuchi, H. (1995). *The knowledge creating company*. New York, NY: Oxford University Press.
- Norman, D. (2002). *The design of everyday things*. New York: Basic Books.
- Polanyi, M. (1967). *The tacit dimension*. London: Routledge & Kegan Paul.
- Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press.
- Stahl, G. (2009). *Studying virtual math teams*. New York, NY: Springer.
- Strayer, D., Drews, F., & Crouch, D. (2006). A comparison of the cell phone driver and the drunk driver. *Human Factors*, 48, 381-391.
- Tajfel, H. (1982). *Social Identity and Intergroup Relations*. London: Cambridge University Press.
- Turner, J. C., Brown, R. J., & Tajfel, H. (1979). Social Comparison and Group Interest in Ingroup Favouritism. *European Journal of Social Psychology*, 9, 187-204.
- Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. New York, NY: Cambridge University Press.
- Wenger, E. (1998). *Communities of Practice: Learning, Meaning and Identity*. New York: Cambridge University Press.
- Wilson, A. (2007). Toward releasing the metadata bottleneck: a baseline evaluation of contributor-supplied metadata. *Library Resources and Technical Services*, vol. 51, No. 1. (January 2007), pp. 16-28.
- Zhou, X.; Hu, X.; Zhang, X.; Lin, X.; & Song, I. (2006). Context-sensitive semantic smoothing for the language modeling approach to genomic IR. ACM SIGIR 2006 (Aug 6-11, 2006, Seattle, WA, USA), 170-177.
- Zhou, X., Hu, X., Zhang, X., (2007a). Topic signature language models for ad hoc retrieval. *IEEE Transactions on Knowledge and Data Engineering*. 19(9), 1276-1287.

## Michael Khoo

Assistant Professor, College of Information Science and Technology, Drexel University  
tel: 215.895.1230 (office) email: Michael.khoo@ischool.drexel.edu  
web: <http://www.ischool.drexel.edu/faculty/mkhoo/>

**Michael Khoo's** research investigates the socio-technical dimensions of technology use, with a focus on information systems and digital libraries, and understanding the different understandings and practices that users bring to their interactions with these systems. He draws on models of culture, practice and knowledge, from anthropology, communication studies, user-centered design, and sociotechnical studies, and combines these with an extensive knowledge of digital library operations and cataloging practices. He has worked with a range of digital libraries, including three NSF-funded projects (the Digital Water Education Library, the Digital Library for Earth System Education, and the National Science Digital Library). He currently conducts research with the Internet Public Library.

### Professional Preparation

Bulmershe College, University of Reading, UK	Human Geography	B.A. (Honors), 1983
School of Advanced Study, University of London, UK	Area Studies	M.A., 1984
University of Colorado, USA	Cultural Anthropology	M.A, 2000
University of Colorado, USA	Communication	Ph.D., 2004
National Center for Atmospheric Research, Advanced Study Program	Digital libraries	2004-2005

### Appointments and Professional Experience

2007-	Assistant Professor	College of Information Science and Technology Drexel University, Philadelphia, USA
2005-07	Evaluation Coordinator	National Science Digital Library, University Corporation for Atmospheric Research, Boulder, Colorado, USA

### Recent Relevant Publications

- 2009** **Khoo, M., X. Lin, & J. Park.** *A User-Friendly Metadata Quality Control Tool for the Internet Public Library.* 9th ACM/IEEE Joint Conference on Digital Libraries (JCDL), Austin, TX, June 15-19, 2009, pp. 407-408.
- 2008** **Khoo, M., J. Pagano, A. Washington, M. Recker, B. Palmer, & R. Donahue.** Using Web Metrics to Analyze Digital Libraries. Joint Conference on Digital Libraries (JCDL), Pittsburgh, PA, June 16-20, 2008. Pp. 375-384.
- 2007** **Khoo, M.** Guest editor, Special issue of the New Review of Hypermedia and Multimedia: 'Studying the Users of Digital Educational Technologies,' Vol. 13(2).
- 2006** **Khoo, M.** A Sociotechnical Framework for Digital Library Evaluation. Procs. Annual Meeting of the American Society for Information Science and Technology (ASIST), Austin, TX, November 3-6, 2006.
- 2005** **Khoo, M.** The Tacit Dimensions of User Behavior: The Case of the Digital Water Education Library. Joint Conference on Digital Libraries (JCDL), Denver, CO, June 7-11, 2005. Pp. 213-222.

### Five Other Significant Publications

- 2009** **Khoo, M., & S. Giersch.** *Planning an Evaluation Initiative.* In: Papatheodorou, C., & G. Tsakonias (Eds.), *Evaluating Digital Libraries*, Oxford, U.K.: Chandos Publishing House.
- 2004** **Khoo, M.** The Use of Visual Artifacts in the User-Centred Design of Educational Digital Libraries. *Journal of Digital Information*, Volume 5, Issue 3.

- 2003 Sumner, T., M. Khoo, and M. Recker.** Understanding Educator Perceptions of “Quality” in Digital Libraries. Joint Conference on Digital Libraries (JCDL), Houston, TX, May 27-31, 2003.
- 2002 Khoo, M., H. Devaul, and T. Sumner.** Functional Requirements for Groupware to Support Community-Led Collections Building. European Conference on Digital Libraries (ECDL), Rome, September 16-18, 2002. Pp. 190-3.
- 2001 Khoo, M.** Ethnography, Evaluation, and Design as Integrated Strategies: A Case Study from WES. European Conference on Digital Libraries (ECDL), Darmstadt, Germany, September 4-8, 2001. Pp. 263-274.

### **Synergistic Activities**

- 2008: OCLC Library and Information Science Research Grant Program: “Addressing the “metadata bottleneck” by developing and evaluating an online tool to support non-specialists to evaluate dublin core metadata records.” \$14,855.
- 2008-present: Program committee member, European Conference on Digital Libraries (ECDL)
- 2008-present: Reviewer, Journal of the American Society for Information Science and Technology (JASIST)
- 2005-present: Program committee member, ACM/IEEE Joint Conference on Digital Libraries (JCDL)

### **Collaborators – past 48 months**

Jean-Françoise Blanchette	UCLA
Geof Bowker	Santa Clara University
George Buchanan	Center for HCI Design, City University, London, UK
Sally Jo Cunningham	Dept. Computer Science, Waikato University, New Zealand
Robert Donahue	Teachers’ Domain Digital Library, WGBH, Boston
Luis Francisco-Revilla	University of Texas – Austin
Gary Geisler	University of Texas – Austin
David Germano	Dept. Religious Studies, University of Virginia
Sarah Giersch	Center for Digital Research and Scholarship, Columbia University
Xia Lin	The iSchool, Drexel University
Rebecca Menendez	Autry Museum, Los Angeles
David MacArthur	School of Education, University of North Carolina
Joe Pagano	The Library of Congress
Bart Palmer	Utah State University
Jung-ran Park	The iSchool, Drexel University
Mimi Recker	Utah State University
Anne Washington	The Library of Congress
Lynn Westbrook	University of Texas - Austin
Lee Zia	Division of Undergraduate Education, National Science Foundation

### **Graduate Advisors**

- Michele H. Jackson, Dept. Communication, University of Colorado-Boulder
- Tamara (“Tammy”) Sumner, Depts. Cognitive and Computer Science, University of Colorado-Boulder

### **Thesis Advisor and Postgraduate-Scholar Sponsor**

None.

## **Eileen G. Abels**

Professor, College of Information Science and Technology  
College of Information Science & Technology, Drexel University  
Email: [eabels@drexel.edu](mailto:eabels@drexel.edu) Phone: 215-895-6274 (o)

### **Professional Preparation**

Ph.D., Library & Information Science, University of California Los Angeles	1985
Master's in Information and Library Science, University of Maryland	1977
Bachelor's in French, Clark University	1974

### **Professional Positions**

Jan. 2007 -	College of Information Science and Technology, Drexel University Professor and Associate Dean (Previously Associate Professor or & Program Director)
Aug. 1991- Dec 2006	College of Information Studies, University of Maryland Assistant professor 1991 – 1997; Associate Professor 1997 – 2006.
Sept. 1989 - Aug. 1991	Price Waterhouse - Washington National Tax Service, Washington, D.C. Librarian
Feb. 1986 - April 1988	Boston Consulting Group, Los Angeles, California West Coast Business Information Specialist

### **Honors and Awards**

- ISI Outstanding Information Science Teacher Award 2008.
- SLA's Rose L. Vormelker Award 2007.
- Winner ALISE Research Award Competition 2007. [With Denise Agosto and Lorri Mon]
- Ida and George Eliot Prize. 2004. With Keith Cogdill and Lisl Zach
- SLA's Factiva Leadership Award. 2004.

### **Publications**

#### **Five Relevant Publications**

1. Denise E. Agosto, Eileen G. Abels, Lorri Mon, Lydia Harris. (2009). The Internet Public Library as Service-Based Learning. IN: Service Learning. Edited by Lorie Roy, Alex Hershey Meyers, and Kelly Jensen. Chicago: American Library Association.
2. Mon, Lorri; Abels, Eileen G.; Agosto, Denise E.; Japzon, Andrea; Most, Linda; Masnik, Mike; Hamann, Jeanne. (Summer 2008). Remote Reference in U.S. Public Library Practice and LIS Education. *Journal of Education for Library and Information Science*, Vol. 49(3): 180-194.
3. Abels, Eileen G.; White, Marilyn Domas; Kim, Soojung. (Jan 2007) Developing Subject-related Web Sites Collaboratively: The Virtual Business Information Center. *Journal of Academic Librarianship* 33(1): 27-40.
4. Abels, Eileen G. and Ruffner, Malissa. (2005). Online Virtual Reference Training. IN: *The Virtual Reference Desk: Creating a Reference Future*. Edited by R. David Lankes, Eileen G. Abels, Marilyn White and Saira N. Haque. New York: Neal Schuman Publishers, Inc.

5. Abels, Eileen. Information Seekers' Perspectives of Libraries and Librarians. (2004) IN: Nitecki, Danuta, ed. *Advances in Librarianship*. Vol. 28, pps 151-169 Amsterdam: Elsevier.

### **Five Other Publications**

1. Lankes, David R.; Abels, Eileen G.; White, Marilyn Domas; Haque, Saira N., eds. (2006). *The Virtual Reference Desk: Creating a Reference Future* Neal Shuman Publishers.
2. Lin, Jimmy; Wu, Philip; Demner-Fushman, Dina; Abels, Eileen G. (August 2006) Exploring the Limits of Single-Iteration Clarification Dialogs. *29th Annual International ACM SIGIR Conference on Research & Development on Information Retrieval*. Seattle, Washington. pp. 469-476.
3. Abels, Eileen G., Cogdill, Keith, and Zach, Lisl. (2002) The Contributions of Library and Information Services to Hospitals and Academic Health Sciences Centers: A Preliminary Taxonomy. *Journal of the Medical Library Association*. 90(3):276-283.
4. Abels, Eileen G. and Magi, Trina. (2001) Current Practices and Trends in 20 Top Business School Libraries. *Journal of Business and Finance Librarianship*, 6(3).
5. White, Marilyn Domas; Abels, Eileen G.; and Gordon-Murnane, Laura. (1998) What Constitutes Adoption of the Web: A Methodological Problem in Assessing Adoption of the World Wide Web for Electronic Commerce. *ASIS 98 Information Access in the Global Information Economy*. October 25-29, 1998. Medford, NJ: Information Today, Inc. Volume 35: 217-224.

### **Synergistic Activities**

- 7/07 Laura Bush 21<sup>st</sup> Century Librarian Program. A Virtual Learning Laboratory for Digital Reference: Transforming the Internet Public Library. Principal Investigator \$613,478. Denise Agosto, Co-PI; partners Florida State University and University of Michigan.
- 9/05 CLIMB 2 at UMD. Judith Klavans, P.I. and Eileen Abels, Senior Personnel. \$841,000 September 2005 – August 2007. Other investigators: Jimmy Lin and Dagobert Soergel.
- 1/04 BRIDGE Technical Task Order 31 Enhancements to BRDGE, a Bilingual Inference and Dictionary Generation Environment. July 1, 2004 – December 31, 2005. \$677,590. Funding Agency Maryland Procurement. Co-PI. PI -- Dave Doermann.
- 10/00 Medical Library Association grant with Keith Cogdill (PI), Eileen Abels (Co-PI), and Lisl Zach (doctoral student) -- Measuring and Communicating the Value of Information Services. \$49,580. Accepted March 31, 2000. Research began June 2000 (duration 18 months).
- 9/99 Contract with Montgomery County to continue the development and maintenance of MC Info in collaboration with the Department of Economic Development and the Montgomery County Public Libraries. \$35,400.
- 7/97 Economic and Community Survey in Collaboration with a County Public Library in Maryland. Kellogg Foundation Grant for \$100,000 with Paul Wasserman and Gary Marchionini. Co-PI. Renewal 7/98 \$100,000. Principal Investigator.
- 3/93 Factors influencing adoption and use of electronic networks among scientists and engineers at small universities and colleges. Principal Investigator. National Science Foundation. \$30,000. Co-investigators Peter Liebscher and Diane Barlow.

### **Collaborators**

Denise Agosto (Drexel University), Lorri Mon (Florida State University)

**Graduate Advisor:** Harold Borko (UCLA)

**Sean Goggins** College of Information Science and Technology  
 Drexel University Philadelphia, PA 19104  
 215-948-2729 (office)  
 215-895-2494 (fax)  
 sgoggins@drexel.edu

**Sean Goggins** teaches, publishes and conducts research in the uptake and use of information and communication technologies by distributed teams of students and workers. Dr. Goggins’ research interests are centered on socio-technical systems and theory to support physically distributed small group collaboration and creativity. He pursues this agenda through design based research methods, quantitative analysis of joint performance outcomes among small groups and deep ethnographic studies of small online group formation and development. Dr. Goggins obtained his graduate education while working in technology leadership positions in the utility, software, medical device and publishing industries.

**Professional Preparation**

University of Wisconsin Madison	History	BS 1989
University of Wisconsin Milwaukee	Human Resource Development	MS 1990
University of Minnesota Twin Cities	Software Engineering (Computer Science)	MS 2003
University of Missouri Columbia	Information Science and Learning Technology	PhD 2009

**Appointments & Professional Experience**

- 2009-present **Assistant Professor**  
 College of Information Science & Technology  
 Drexel University, Philadelphia, PA
- 2009 **Teaching Assistant**  
 University of Missouri
- 2005-2008 **Director of Product Development**  
 Foliotek  
 Columbia, MO
- 2004-2005 **Software Architect**  
 MDConsult / Elsevier, St. Louis, MO
- 2003-2004 **Software Architect**  
 Nestle Purina, St. Louis, MO
- 1999-2003 **Enterprise Architect**  
 Guidant Corporation, St. Paul, MN

**Relevant Publications**

Tsai, I.-C., Kim, B., Goggins, S., Kumalasari, C., Laffey, J., & Amelung, C. (2008). “A Model Explaining the Social Nature of Online Learning,” *Journal of Educational Technology and Society*.  
 Laffey, J., Amelung, C., & Goggins, S. (2008). “A Context Awareness System for Online Learning: Design Based Research,” *International Journal on E-Learning*.



- Schmidt, M., Goggins, S., & Laffey, J. (2008). "The Design of iSocial," *International Journal of Technology, Knowledge and Society*.
- Goggins, S., Laffey, J., & Tsai, I.-C. (2007). "Cooperation and Groupness: Community Formation in Small online Collaborative Groups," *Paper presented at the ACM Group, 2007*, Sanibel Island, FL.
- Goggins, S., Tsai, I.-C., Kim, B., Kumalasari, C., Laffey, J., & Amelung, C. (2007). "Building a Model Explaining the Social Nature of Online Learning," *Paper presented at the American Education Research Association, 2007*, Chicago, IL.

### **Other Publications**

- Goggins, S., & Erdelez, S. (2009). *HIB and HCI: Common Interests in Different Communities*. Paper presented at the iConference, 2009, Chapel Hill, NC.
- Goggins, S., & Erdelez, S. (2009). Collaborative Information Behavior in Completely Online Groups. In J. Foster (Ed.), *Collaborative Information Behavior: User Engagement and Communication Sharing*. Hershey, PA: ISI Global.
- Goggins, S., Floyd, I., Sawyer, S., Grudin, J., Dabbish, L., Erickson, I. et al. (2009). The Science of Socio-Technical Systems in iSchools. *iConference, 2009*.
- Goggins, S., Laffey, J., & Galyen, K. (2009). *Social Ability in Online Groups: Representing the Quality of Interactions in Social Computing Environments*. Paper presented at the IEEE Social Computing Conference, Vancouver, BC.
- Laffey, J., Hong, R.-Y., Galyen, K., Goggins, S., & Amelung, C. (2009). *Context Aware Activity Notification System: Supporting CSCL*. Poster presented at the CSCL, 2009, Rhodes, Greece.
- Stichter, J., Schmidt, C., Schmidt, M., Goggins, S., Babuich, R., & Laffey, J. (2009). *iSocial: A 3-D Virtual Learning Environment for Enhanced Social Interaction and Development of Social Competence*. Poster presented at the CSCL, 2009, Rhodes, Greece.

### **Collaborators & Other Affiliations**

*Collaborators and Co-Editors:* Sanda Erdelez (Missouri), James Laffey (Missouri), Chris Amelung (Yale), Ichun Tsai (University of Akron), Gerry Stahl (Drexel), Carolyn Rose (CMU), Matt Schmidt (Missouri).

*Dissertation Advisors:* James Laffey, Sanda Erdelez, Chi-Ren Shyu & Joi Moore (University of Missouri – Columbia).

## Biographical Sketch

Jiexun Li  
Assistant Professor  
College of Information Science and Technology , Drexel University  
3141 Chestnut Street, Philadelphia, PA 19104  
215-895-1459  
Jiexun.Li@drexel.edu

### A. PROFESSIONAL PREPARATION

<u>College/University</u>	<u>Major</u>	<u>Degree &amp; Year</u>
Tsinghua University, China	Management Information Systems	Bachelor, 2000
Tsinghua University, China	Management Science	Master, 2002
University of Arizona	Management Information Systems	Ph.D., 2007

### B. ACADEMIC/PROFESSIONAL APPOINTMENTS

Assistant Professor 2007~present  
College of Information Science and Technology, Drexel University, Philadelphia, PA

Research Associate 2002~2007  
Artificial Intelligence Lab, Department of Management Information Systems,  
Eller College of Management, University of Arizona, Tucson, AZ

### C. PUBLICATIONS

#### Publications Most Closely Related to Proposal

- Lin, X., Li, J., & Zhou, X. (2008). Theme creation for digital collections. *Proceedings of the International Dublin Core and Metadata Applications (DC 2008)*, 34-42, September 22-26, Berlin, German.
- Li, J., Zhang, Z., Li, X., & Chen, H. (2008). Kernel-based learning for biomedical relation extraction. *Journal of the American Society for Information Science and Technology*, 59(5), 756-769.
- Li, J., Zheng, R., & Chen, H., (2006). From fingerprint to writeprint. *Communications of the ACM*, 49(4), 76-82.
- Zheng, R., Li, J., Chen, H., Huang, Z., & Qin, Y. (2006). A framework of authorship identification for online messages: Writing style features and classification techniques," *Journal of the American Society for Information Science and Technology*, 57(3), 378-393.

#### Other Significant Publications

- Li, X., Chen, H., Zhang, Z., Li, J., & Nunamaker, J. F. (2009). Managing knowledge in light of its evolution process: An empirical study on citation network-based patent classification, *Journal of Management Information Systems*, 26(1), 129-153.
- Li, J., Wang, H. J., Zhang, Z., & Zhao, J. L. (2009). A policy-based process mining framework: Mining business policy texts for discovering process models, *Information Systems and e-Business Management*, (published online: April 11, 2009).
- Li, J., Wang, H. J., Zhang, Z., & Zhao, J. L. (2008). Relation-centric task identification for policy-based process mining. *Proceedings of the International Conference on Information Systems (ICIS 2008)*, December 14-17, Paris, France.
- Xu, J., Wang, A. G., Li, J., & Chau, M. (2007). Complex problem solving: A case study on identity matching based on social contextual information. *Journal of the Association for Information Systems: Special Issue Enid Mumford's Contribution to Information Systems Theory and Theoretical Thinking*, 8(10), Article 31.

- Li, J., Su, H., Chen, H., & Futscher, B. W. (2007). Optimal search-based gene subset selection for gene array cancer classification. *IEEE Transactions on Information Technology in Biomedicine*, 11(4), 398-405.
- Huang, Z., Li, J., Su, H., Watts, G. S., & Chen, H. (2007). Large-scale regulatory network analysis from microarray data: Modified Bayesian network learning and association rule mining, *Decision Support Systems*, 43, 1207-1225.
- Li, J., Li, X., Su, H., Chen, H., & Galbraith, D.W. (2006). A framework of integrating gene functional relations from heterogeneous data sources: An experiment on *Arabidopsis thaliana*. *Bioinformatics*, 22(16), 2037-2043.

#### **D. SYNERGISTIC ACTIVITIES**

- Discourse Analysis of the Question-Answering Service:  
I investigated the linguistic stylometric patterns in online question-answering communications between patrons and librarians in digital libraries. This study is aimed at enhance the effectiveness of interactions and collaborations during utilization of digital information services.
- A Probabilistic Relational Model (PRM) Based Approach for Identity Matching:  
I developed a novel approach to derive personal and social identity features based on a relational database schema. The derived social features that represent the characteristics of people’s social activities and relations can help tackle identity management problems, e.g., identity deception detection in crime investigation.
- A Framework of Integrating Biomedical Relations from Heterogeneous Data Sources:  
I developed a Bayesian framework that can combine biomedical relations mined from various data sources into an integrated network. This integrated network was shown to be more reliable and can help biomedical researchers generate new hypotheses and potentially lead to new findings. This framework can be applicable to other similar knowledge integration problems as well.

#### **E. COLLABORATORS AND OTHER AFFILIATIONS**

##### **Collaborators over the Last 48 Months:**

Neal Handly, M.D. (College of Medicine, Drexel University):  
 Prediction of Patients to be admitted through the Emergency Department 2008 ~ present

Hsinchun Chen, Ph.D. (Department of Management Information Systems, University of Arizona):  
 #0429364 Chen & Atabakhsh (PI's) 10/01/2004 – 09/30/2007  
 NSF/Digital Government (GD) Program  
 COPLINK Center: Social Network Analysis and Identity Deception Detection for Law Enforcement and Homeland Security

##### **Graduate Advisor**

Hsinchun Chen, Ph.D. (Department of Management Information Systems, University of Arizona)

## Gerry Stahl

College of Information Science  
and Technology  
Drexel University  
Philadelphia, PA 19104

215-895-0544 (office)  
215-895-2494 (fax)  
gerry.stahl@drexel.edu  
www.cis.drexel.edu/faculty/gerry

**Gerry Stahl** teaches, publishes and conducts research in human-computer interaction (HCI) and computer-supported collaborative learning (CSCL). His books are *Group Cognition: Computer Support for Building Collaborative Knowledge* (2006, MIT Press) and *Studying Virtual Math Teams* (2009, Springer). He is founding Executive Editor of the *International Journal of Computer-Supported Collaborative Learning (ijCSCL)*. He is the Principal Investigator of the *Virtual Math Teams Project*, a large 6-year research effort in collaboration with the Math Forum@Drexel. He served as Program Chair for the international CSCL '02 conference and Workshops Chair for CSCL '03, '05, '07 and '09. He teaches undergraduate, masters and PhD courses in HCI, CSCW and CSCL at the I-School of Drexel.

### Professional Preparation

Massachusetts Institute of Technology (MIT)	Humanities & Science (Math & Philosophy)	BS 1967
University of Heidelberg	Continental Philosophy	1967-68
University of Frankfurt	Social Theory	1971-73
Northwestern University	Philosophy	MA 1971
Northwestern University	Philosophy	PhD 1975
University of Colorado	Computer Science	MS 1990
University of Colorado	Computer Science	PhD 1993
University of Colorado	Computer Science & Cognitive Science	Postdoc 1996-99

### Appointments & Professional Experience

2002-present	<b>Associate Professor</b> (Tenured May 2008) College of Information Science & Technology Drexel University, Philadelphia, PA
2001-2002	<b>Visiting Research Scientist</b> BSCW Development Team, CSCW Department, FIT GMD and Fraunhofer Institutes, Bonn, Germany
1999-2001	<b>Assistant Research Professor</b> Department of Computer Science & Institute of Cognitive Science University of Colorado, Boulder, CO
1996-1999	<b>Post Doctoral Research Fellow</b> Center for LifeLong Learning and Design University of Colorado, Boulder, CO
1993-1996	<b>Director of Software R&amp;D</b> Owen Research Inc., Boulder, CO

### Relevant Publications

- Stahl, G. (2006). Group cognition in an online chat community: Analyzing collaborative use of a cognitive tool. *Journal of Educational Computing Research (JECR) special issue on Cognitive tools for collaborative communities*. Available online at <http://www.cis.drexel.edu/faculty/gerry/pub/jecr.pdf>.
- Stahl, G. (2006). Sustaining group cognition in a math chat environment. *Research and Practice in Technology Enhanced Learning (RPTTEL)*, 1 (2). Available online at <http://www.cis.drexel.edu/faculty/gerry/pub/rptel.pdf>.
- Stahl, G. (2006). Analyzing and designing the group cognitive experience. *International Journal of Cooperative Information Systems (IJCIS)*. Available online at <http://www.cis.drexel.edu/faculty/gerry/pub/ijcis.pdf>.
- Stahl, G. (2005). Group cognition in computer assisted learning. *Journal of Computer Assisted Learning (JCAL)*. Available online at <http://www.cis.drexel.edu/faculty/gerry/publications/journals/JCAL.pdf>.
- Stahl, G., Rohde, M., & Wulf, V. (2006). Introduction: Computer support for learning communities. *Behavior and Information Technology (BIT)*. Available online at [http://www.cis.drexel.edu/faculty/gerry/pub/bit\\_intro.pdf](http://www.cis.drexel.edu/faculty/gerry/pub/bit_intro.pdf).

## Other Publications

- Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press. Available online at <http://www.cis.drexel.edu/faculty/gerry/mit/>.
- Stahl, G. & Hesse, F. (2006). Inaugural issue. *International Journal of Computer-Supported Collaborative Learning (ijCSCL)*, 1 (1). Available online at <http://ijCSCL.org>.
- Stahl, G. (Ed.). (2002). *Computer support for collaborative learning: Foundations for a CSCL community*. Proceedings of CSCL 2002. January 7-11. Boulder, Colorado, USA. Hillsdale, NJ: Lawrence Erlbaum Associates Available online at <http://isls.org/cscl/cscl2002proceedings.pdf>.
- Stahl, G. (2005). *Groups, group cognition & groupware [keynote]*. Paper presented at the International Workshop on Groupware (CRIWG 2005), Racife, Brazil. Available online at <http://www.cis.drexel.edu/faculty/gerry/pub/criwg2005.pdf>.
- Stahl, G. (2003). *The future of computer support for learning: An American/German DeLFIc vision [keynote]*. Paper presented at the First Conference on e-Learning of the German Computer Science Society (DeLFI 2003), Munich, Germany. Proceedings pp. 13-16. Available online at <http://www.cis.drexel.edu/faculty/gerry/publications/presentations/delfi>.

## Synergistic Activities

- 2007-2008: “*Exploring Adaptive Support for Virtual Math Teams.*” (co-PI with PI Carolyn Rose) \$50,000; sponsor: NSF SGER.
- 2005-2007: “*SoL Catalyst: Engaged Learning in Online Communities.*” (PI with co\_PIs Sharon Derry, Mary Marlino, K. Ann Renninger, Daniel Suthers, Stephen Weimar) \$180,762; sponsor: NSF SOL.
- 2003-2008: “*IERI: Catalyzing & Nurturing Online Workgroups to Power Virtual Learning Communities.*” (PI with co-PIs Stephen Weimar and Wesley Shumar) \$2,300,000; sponsor: NSF IERI.
- 2003-2005: “*Collaboration Services for the Math Forum Digital Library*” (PI with co-PIs Stephen Weimar and Wesley Shumar) \$450,000; sponsor: NSF NSDL.
- 1997-2000: “*Allowing Learners to be Articulate: Incorporating Automated Text Evaluation into Collaborative Software Environments*” (primary author and primary software developer; PIs: Gerhard Fischer, Walter Kintsch and Thomas Landauer) \$678,239; sponsor: James S. McDonnell Foundation.
- 1997-2000: “*Conceptual Frameworks and Computational Support for Organizational Memories and Organizational Learning*” (co-PI with Gerhard Fischer and Jonathan Ostwald), \$725,000; sponsor: NSF.
- 1998-1999: “*Collaborative Web-Based Tools for Learning to Integrate Scientific Results into Social Policy*” (co-PI with Ray Habermann) \$89,338; sponsor: NSF.

## Collaborators & Other Affiliations

*Scientific Advisory Boards:* Knowledge Media Research Center (KMRC, Germany), Learning Sciences Laboratory (LSL, NIE, Singapore), Knowledge Practices Laboratory (K-P Lab, Finland).

*Collaborators and Co-Editors:* Clarence (Skip) Ellis, Gerhard Fischer, Raymond Habermann, Walter Kintsch, Thomas Landauer, Curtis LeBaron, Raymond McCall, Jonathan Ostwald, Alexander Repenning, Tamara Sumner (U. Colorado, Boulder); Robert Allen, K. Ann Renninger, Wesley Shumar, Stephen Weimar, Alan Zemel (Drexel U., Philadelphia); Timothy Koschmann (Southern Illinois U.); Angela Carell, Thomas Herrmann, Andrea Kienle, Ralf Klamma, Kai-Uwe Loser, Wolfgang Prinz, Markus Rohde, Volker Wulf (Germany); Sten Ludvigsen, Anders Morch, Barbara Wasson (Norway), Cesar Alberto Collazos (Chile); Jan-Willem Strijbos (Netherlands). Carolyn Rose (CMU), Daniel Suthers (Hawaii), Sharon Derry (Wisconsin), Mary Marlino (UCAR)

*Dissertation Advisors:* Gerhard Fischer, Clayton Lewis, Raymond McCall, Mark Gross (U. Colorado, Boulder). Samuel Todes, Theodor Kiesel (Northwestern).

*Graduate Students, Post-Docs, visiting Researchers:* Rogerio dePaula, Elizabeth Lenell, Alena Sanusi, David Steinhart (U. Colorado, Boulder); Murat Cakir, Ilene Litz Goldman, Trish Grieb-Neff, Yolanda Jones, Wanda Kunkle. Deb LeBelle, Debra McGrath, Pete Miller, Johann Sarmiento, Ramon Toledo, Jim Waters, Alan Zemel, Nan Zhou (Drexel U., Philadelphia); Andrea Kienle (U. Dortmund, Germany); Cesar Alberto Collazos (U. Chile, Chile); Jan-Willem Strijbos (Open U., Netherlands); Fatos Xhafa (Open U. Catalonia, Spain); Stefan Trausan-Matu (Politechnica University of Bucharest, Romania); Angela Carell (Bochum U., Germany); Martin Wesner, Martin Mühlpfordt (FhG-IPSI, Germany); Elizabeth Charles (Canada), Weiquin Chen (Norway).

## Biographical Sketch

### Xia Lin, Ph.D.

Associate Professor  
College of Information Science and Technology  
Drexel University, Philadelphia, PA 19104  
Phone: (215) 895-2482 FAX: (215) 895-2494  
E-mail: xlin@drexel.edu

### Education

1993 PhD in Information Science, University of Maryland.  
1986 MLS in Library & Information Science, Emory University, Atlanta, Georgia.  
1982 BS in Mathematics, Fujian Teachers University, Fuzhou, China

### Academic Experience

July 2003 – current Associate Professor (tenured)  
College Of Information Science and Technology, Drexel University  
July , 1997 – June 2003 Assistant Professor  
College Of Information Science and Technology, Drexel University  
Aug., 1993 - June, 1997 Assistant Professor  
School of Library and Information Science, University of Kentucky

### Research Areas

Information Visualization  
Digital Libraries and Information Retrieval  
Visual interface design  
Thesaurus, Ontology, and Knowledge Mapping

### Related Publications

1. Sofia J. Athenikos, S. & **Lin, X.** (2009). "WikiPhiloSofia: Extraction and Visualization of Facts, Relations, and Networks Concerning Philosophers Using Wikipedia". Paper presented at the 2009 Digital Humanities Conference (DH 2009) (University of Maryland, College Park, MD, USA, 22-25 June 2009). pp. 56 -62.
2. **Lin, X.**; Li, J.; & Zhou, X. (2008). Theme Creation for Digital Collections. Proceedings of DC2008, International Conference on Dublin Core and Metadata Applications. (Berlin, Germany; September 22 -26, 2008).
3. Athenikos, S.; **Lin, X.** (2008). The WikiPhil Portal: Visualizing Meaningful Philosophical Connections. Paper presented at the Chicago Colloquium on Digital Humanities and Computer Science (November 1–3, 2008, the University of Chicago).
4. **Lin, X.**; Bui, Y.; & Zhang, D. (2007). Visualization of Knowledge Structure. In: Proceedings of the 11th International Conference of Information Visualization (IV 2007, July 4-6, 2007; Zurich, Switzerland), pp. 476-481.
5. **Lin, X.**; Aluker, S.; Zhu, W.; & Zhang, F. (2006). "Dynamic Concept Representation through a Visual Concept Explorer." Paper presented at the Ninth Conference of the International Society of Knowledge Organization (ISKO 2006, July 4 – 7, Vienna, Austria).
6. White, H.D., **Lin, X.**, Buzydlowski, J., & Chen, C. M. (2004). User-controlled mapping of significant literatures. Proceedings of National Academy of Sciences. 101: 5297-5302.
7. **Lin, X.** (2004). Information Visualization and Content Representation. Modern Technology of Library and Information Service, 2004(10), pp.3-13.
8. **Lin, X.**; White, H. D.; & Buzydlowski, J. (2003). Real-time author co-citation mapping for online searching. International Journal of Information Processing & Management, 39(5), 689-706.

## Xia Lin (2)

### Other Publications

1. Petushi, S. Marker, J. Zhang, J. Zhu, W. Breen, D, Chen, C. **Lin, X.**, Garcia, F. (2008). A Visual Analytics System for Breast Tumor Evaluation. *Analytical and Quantitative Cytology and Histology*, 30:279-290.
2. Zhou, X.; Hu, X.; Zhang, X.; **Lin, X.**; and Il-Yeol Song, I.(2006). "Context-Sensitive Semantic Smoothing for the Language Modeling Approach to Genomic IR", *ACM SIGIR 2006* (Aug 6-11, 2006, Seattle, WA, USA), 170-177.
3. Hu, X., Lin, T.Y., Song I-Y., **Lin, X.**, Yoo I., Song M.(2006). A Semi-supervised Efficient Learning Approach to Extract Biological Relationships from Web-based Biomedical Digital Library, *International Journal of Web Intelligence and Agent System*, Vol .4, No. 3, 2006.
4. Chen, C., **Lin, X.**, Zhu, W. (2006) Trailblazing through a knowledge space of science: Forward citation expansion in CiteSeer. In Grove, Andrew, Eds., *Proceedings of the 69th Annual Meeting of the American Society for Information Science and Technology (ASIS&T 2006)*, Austin, TX. November 3-8, 2006.
5. Chan, M. L.; Lin, X.; & M. L. Zeng (2000). Structural and multilingual approaches to subject access on the web. *IFLA Journal*, 26(3), 187-197.
6. Lin, X.; & Chan, M. L. (1999). Personalized knowledge organization and access for the web. *Library and Information Science Research*, 21(2), 153-172.
7. Lin, X. (1997). Map displays for information retrieval. *Journal of the American Society for Information Science*, 48(1), 40 – 54.
8. Lin, X., Soergel, D., & Marchionini, G. (1991). A self-organizing semantic map for information retrieval. *Proceedings of the Fourteenth Annual International ACM/SIGIR Conference on Research and Development in Information Retrieval*, pp. 262-269.

### Funded Research Activities

- 2009 – 2010 “Planning Grant: I/UCRC Center for Visual Decision Informatics.”  
Co-PI, \$10,000. NSF
- 2007 - 2010 “Developing Faculty in Digital Librarianship for the 21st Century”.  
PI. \$992, 110. IMLS.
- 2004 – 2007 “Preparing Faculty in Management of Digital Information,”  
PI, \$611, 648, IMLS.
- 2005 – 2006 “Predictive Syndromic Surveillance System (PS3).” U.S. Army Medical Research , CO-PI.  
Co-PI, 346,157. U.S. Amy Medical Research.
- 2003 - 2004: “Knowledge Discovery with Information Visualization,”  
Co-PI. SUR Gant (equipment), IBM.
- 2002 – 2003: “Drexel Digital Museum Project” Sponsored by. CO-PI  
Co-PI. \$232,800. The Barra Foundation.

### Synergistic Activities

- Associate Director of Drexel Digital Museum Project
- Editorial Board of the International Journal of Information Visualization
- Reviewer for SIGIR, JCDL, ASIST and several other visualization and IR/DL conferences
- Prototype development for AuthorLink, ConceptLink, PNASLink, Concept Explorer and many others.

### Persons collaborated within the last 48 months:

Dongming Zhang,	Johns Hopkins University
Jian Qin	Syracuse University
Howard White	Drexel University

### Graduate Advisors:

Gary Marchionini Professor of School of Information and Library Science, University of North Carolina.  
Dagobert Soergel, Professor of School of Information Studies, University of Maryland.

Name:

**Dr. Baba Kofi A. Weusijana**

**Objectives:** Educational Software Research and Development, Higher Education Faculty, or Online Teaching Position

**Citizenship:** USA

**Email:** [kofi@edutek.net](mailto:kofi@edutek.net)

**Web:** <http://www.linkedin.com/in/babakofi>

#### EDUCATION

- **Ph.D., Learning Sciences (Educational Software), 2006, Northwestern University, Evanston, IL**  
Dissertation: A Socratic ASK System: Helping Educators Provide a Web-Based Socratic Tutor for Learners.
- **M.S., General Engineering (Client/Server Computing), 1999, San José State University, San José, CA**  
Masters Thesis: Web-Based Student Assessment Internet Software for University Courses
- **B.S., Computer Science (African Studies Minor), 1997, San José State University, San José, CA**
- **B.S., Mathematics, Division of Natural Sciences, 1992, Dillard University, New Orleans, LA**
- **A.S., Computer Science, Technology & Information Systems, 1990, Foothill College, Los Altos Hills, CA**

#### CURRENT POSITION

**Associate Faculty**

Business and Information Technology,  
Cascadia Community College

Bothell, Washington  
March 2008-Present

Teaching programming courses and conducting research on web-based educational software to facilitate critiquing of students' Java code.

#### SELECTED PROFESSIONAL EXPERIENCE

**Research Associate**

NSF Learning in Informal and Formal Environments (LIFE)  
Science of Learning Center (SLC)  
University of Washington, Stanford University, SRI International

Seattle, Washington  
September 2005–  
December 2008

Researched learning environments in Multi-User Virtual Environments (MUVES) and Web 2.0 learning systems. Designed and built Java-based middleware and an intelligent tutoring system client for the Second Life MUVE. Investigated the use and design of educational software for homeschoolers. Fostering collaborative educational outreach and research partnerships, including a partnership with the National Society of Black Engineers (NSBE).

**Technical Co-op**

Adaptive Simulations,  
IBM T. J. Watson Research Center

Yorktown Heights, New  
York  
Summer 2005

Worked with a team of researchers on an authoring tool for a schema-based intelligent tutoring system using Python. Developed and evaluated design wireframes. This work was part of an Adventurous Research project called PASTEL (Pattern-driven Adaptive Simulations for Training, Education, and Learning).

**Consultant**

Academic Technologies,  
Northwestern University

Evanston, Illinois  
February 2003-June 2005

Worked with faculty and a team of developers in the Distributed Learning Group providing Internet tools for instructional use. Developed an Engineering Statistics module using Flash/ActionScript and a component of a content management system using Java and JavaScript. Worked on an inquiry based tutoring system using Flash, Python, and Java. Researched the impact of the ProjectPad Web-Based educational software for the Spoken Word Project.

**Software Engineer**

Center for Connected Learning and Complex  
Modeling (CCL),  
Northwestern University

Evanston, Illinois  
January-July 2001

Worked with a team of programmers and students on the development of NetLogo, a Java educational software application for exploring emergent and complex phenomena with the Logo language. Wrote and maintained the installer, added features, and debugged problems.

**Software Engineer**

AvantSoft Corporation

Sunnyvale, California  
December 1998-  
September 1999

Wrote Java programs and course materials designed to educate customers in the subjects of advanced Java programming, RMI, CORBA, Object Oriented Design, IOP, IDL, Java Drag & Drop, Java Imaging, XML, XML Style Sheets, etc.

**Whitebox SQA Engineer**

NetObjects

Redwood City, California  
August-December 1998

White box testing of Java code and Java Bean components of NetObjects Fusion. Designed and implemented a Java-based test harness. Reviewed components from contractors and 3rd-parties.

**SQA Student Intern**

Symantec Corporation

Cupertino, California  
November 1996-April  
1998

Worked closely with engineers to assure high quality of Symantec's Internet development tools. White box testing of Java Bean components of Visual Café for Java. Wrote Java test suites. Black box & UI testing. Answered email and Usenet questions from customers.



# SUMMARY PROPOSAL BUDGET

YEAR 1

ORGANIZATION <b>Drexel University</b>				FOR NSF USE ONLY			
				PROPOSAL NO.	DURATION (months)		
PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR <b>Michael J Khoo</b>				AWARD NO.	Proposed	Granted	
A. SENIOR PERSONNEL: PI/PI, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)				NSF Funded Person-months		Funds Requested By proposer	Funds granted by NSF (if different)
				CAL	ACAD	SUMR	
1.	<b>Michael J Khoo - PI</b>			0.00	0.00	1.00	\$ <b>10,263</b> \$
2.	<b>Sean Goggins - Co-PI</b>			0.00	0.00	1.00	<b>10,222</b>
3.	<b>Jiexun Li - Co-PI</b>			0.00	0.00	1.00	<b>10,111</b>
4.	<b>Xia Lin - Senior personnel</b>			0.00	0.00	0.00	<b>6,032</b>
5.	<b>Gerry Stahl - Co-PI</b>			0.00	0.00	0.50	<b>6,517</b>
6.	( 0 ) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)			0.00	0.00	0.00	<b>0</b>
7.	( 5 ) TOTAL SENIOR PERSONNEL (1 - 6)			0.00	0.00	3.50	<b>43,145</b>
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)							
1.	( 0 ) POST DOCTORAL SCHOLARS			0.00	0.00	0.00	<b>0</b>
2.	( 1 ) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)			4.00	0.00	0.00	<b>19,800</b>
3.	( 2 ) GRADUATE STUDENTS						<b>50,000</b>
4.	( 0 ) UNDERGRADUATE STUDENTS						<b>0</b>
5.	( 0 ) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)						<b>0</b>
6.	( 0 ) OTHER						<b>0</b>
TOTAL SALARIES AND WAGES (A + B)							<b>112,945</b>
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)							<b>20,269</b>
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)							<b>133,214</b>
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING \$5,000.)							
TOTAL EQUIPMENT							<b>0</b>
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)							<b>9,600</b>
2. FOREIGN							<b>0</b>
F. PARTICIPANT SUPPORT COSTS							
1.	STIPENDS	\$	<u>                    </u>				<b>0</b>
2.	TRAVEL		<u>                    </u>				<b>0</b>
3.	SUBSISTENCE		<u>                    </u>				<b>0</b>
4.	OTHER		<u>                    </u>				<b>0</b>
TOTAL NUMBER OF PARTICIPANTS ( 0 )				TOTAL PARTICIPANT COSTS			<b>0</b>
G. OTHER DIRECT COSTS							
1.	MATERIALS AND SUPPLIES						<b>0</b>
2.	PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION						<b>0</b>
3.	CONSULTANT SERVICES						<b>0</b>
4.	COMPUTER SERVICES						<b>0</b>
5.	SUBAWARDS						<b>0</b>
6.	OTHER						<b>24,600</b>
TOTAL OTHER DIRECT COSTS							<b>24,600</b>
H. TOTAL DIRECT COSTS (A THROUGH G)							<b>167,414</b>
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE) <b>Modified total direct cost (Rate: 53.5000, Base: 144415)</b>							
TOTAL INDIRECT COSTS (F&A)							<b>77,262</b>
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)							<b>244,676</b>
K. RESIDUAL FUNDS							<b>0</b>
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)							\$ <b>244,676</b> \$
M. COST SHARING PROPOSED LEVEL \$ <b>0</b>				AGREED LEVEL IF DIFFERENT \$			
PI/PI NAME <b>Michael J Khoo</b>				FOR NSF USE ONLY			
ORG. REP. NAME* <b>Margaret vigiolto</b>				INDIRECT COST RATE VERIFICATION			
		Date Checked	Date Of Rate Sheet			Initials - ORG	

# SUMMARY PROPOSAL BUDGET

YEAR **2**

ORGANIZATION <b>Drexel University</b>				FOR NSF USE ONLY			
				PROPOSAL NO.	DURATION (months)		
PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR <b>Michael J Khoo</b>				AWARD NO.	Proposed	Granted	
A. SENIOR PERSONNEL: PI/PI, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)				NSF Funded Person-months		Funds Requested By proposer	Funds granted by NSF (if different)
				CAL	ACAD	SUMR	
1.	<b>Michael J Khoo - PI</b>			0.00	0.00	0.00	\$ <b>10,777</b>
2.	<b>Eileen Abels - Co-PI</b>			0.00	0.00	0.50	<b>7,229</b>
3.	<b>Sean Goggins - Co-PI</b>			0.00	0.00	1.00	<b>10,734</b>
4.	<b>Jiexun Li - Co-PI</b>			0.00	0.00	0.00	<b>10,617</b>
5.							
6.	( <b>0</b> ) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)			0.00	0.00	0.00	<b>0</b>
7.	( <b>4</b> ) TOTAL SENIOR PERSONNEL (1 - 6)			0.00	0.00	1.50	<b>39,357</b>
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)							
1.	( <b>0</b> ) POST DOCTORAL SCHOLARS			0.00	0.00	0.00	<b>0</b>
2.	( <b>1</b> ) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)			4.00	0.00	0.00	<b>20,790</b>
3.	( <b>2</b> ) GRADUATE STUDENTS						<b>52,500</b>
4.	( <b>0</b> ) UNDERGRADUATE STUDENTS						<b>0</b>
5.	( <b>0</b> ) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)						<b>0</b>
6.	( <b>0</b> ) OTHER						<b>0</b>
TOTAL SALARIES AND WAGES (A + B)							<b>112,647</b>
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)							<b>19,368</b>
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)							<b>132,015</b>
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING \$5,000.)							
TOTAL EQUIPMENT							<b>0</b>
E. TRAVEL							<b>10,080</b>
1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)							<b>10,080</b>
2. FOREIGN							<b>0</b>
F. PARTICIPANT SUPPORT COSTS							
1.	STIPENDS	\$	<b>0</b>				
2.	TRAVEL		<b>0</b>				
3.	SUBSISTENCE		<b>0</b>				
4.	OTHER		<b>0</b>				
TOTAL NUMBER OF PARTICIPANTS ( <b>0</b> )							
TOTAL PARTICIPANT COSTS							<b>0</b>
G. OTHER DIRECT COSTS							
1. MATERIALS AND SUPPLIES							<b>0</b>
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION							<b>0</b>
3. CONSULTANT SERVICES							<b>0</b>
4. COMPUTER SERVICES							<b>0</b>
5. SUBAWARDS							<b>0</b>
6. OTHER							<b>25,830</b>
TOTAL OTHER DIRECT COSTS							<b>25,830</b>
H. TOTAL DIRECT COSTS (A THROUGH G)							<b>167,925</b>
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE) <b>Modified total direct cost (Rate: 54.2500, Base: 143775)</b>							
TOTAL INDIRECT COSTS (F&A)							<b>77,998</b>
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)							<b>245,923</b>
K. RESIDUAL FUNDS							<b>0</b>
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)							\$ <b>245,923</b> \$
M. COST SHARING PROPOSED LEVEL \$ <b>0</b>				AGREED LEVEL IF DIFFERENT \$			
PI/PI NAME <b>Michael J Khoo</b>				FOR NSF USE ONLY			
ORG. REP. NAME* <b>Margaret vigiolto</b>				INDIRECT COST RATE VERIFICATION			
		Date Checked	Date Of Rate Sheet	Initials - ORG			

# SUMMARY PROPOSAL BUDGET

YEAR 3

ORGANIZATION <b>Drexel University</b>				FOR NSF USE ONLY			
				PROPOSAL NO.	DURATION (months)		
PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR <b>Michael J Khoo</b>				AWARD NO.	Proposed	Granted	
A. SENIOR PERSONNEL: PI/PI, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)				NSF Funded Person-months		Funds Requested By proposer	Funds granted by NSF (if different)
				CAL	ACAD	SUMR	
1.	<b>Michael J Khoo - PI</b>			0.00	0.00	1.00	\$ <b>11,316</b> \$
2.	<b>Sean Goggins - Co-PI</b>			0.00	0.00	1.00	<b>11,271</b>
3.	<b>Jiexun Li - Co-PI</b>			0.00	0.00	1.00	<b>11,148</b>
4.							
5.							
6.	( 0 ) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)			0.00	0.00	0.00	<b>0</b>
7.	( 3 ) TOTAL SENIOR PERSONNEL (1 - 6)			0.00	0.00	3.00	<b>33,735</b>
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)							
1.	( 0 ) POST DOCTORAL SCHOLARS			0.00	0.00	0.00	<b>0</b>
2.	( 1 ) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)			4.00	0.00	0.00	<b>21,830</b>
3.	( 2 ) GRADUATE STUDENTS						<b>55,125</b>
4.	( 0 ) UNDERGRADUATE STUDENTS						<b>0</b>
5.	( 0 ) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)						<b>0</b>
6.	( 0 ) OTHER						<b>0</b>
TOTAL SALARIES AND WAGES (A + B)							<b>110,690</b>
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)							<b>17,892</b>
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)							<b>128,582</b>
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING \$5,000.)							
TOTAL EQUIPMENT							<b>0</b>
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)							<b>10,584</b>
2. FOREIGN							<b>9,000</b>
F. PARTICIPANT SUPPORT COSTS							
1.	STIPENDS	\$	<b>0</b>				
2.	TRAVEL		<b>0</b>				
3.	SUBSISTENCE		<b>0</b>				
4.	OTHER		<b>0</b>				
TOTAL NUMBER OF PARTICIPANTS ( 0 ) TOTAL PARTICIPANT COSTS							<b>0</b>
G. OTHER DIRECT COSTS							
1.	MATERIALS AND SUPPLIES						<b>0</b>
2.	PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION						<b>0</b>
3.	CONSULTANT SERVICES						<b>0</b>
4.	COMPUTER SERVICES						<b>0</b>
5.	SUBAWARDS						<b>0</b>
6.	OTHER						<b>27,122</b>
TOTAL OTHER DIRECT COSTS							<b>27,122</b>
H. TOTAL DIRECT COSTS (A THROUGH G)							<b>175,288</b>
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE) <b>Modified total direct cost (Rate: 54.5000, Base: 149930)</b>							
TOTAL INDIRECT COSTS (F&A)							<b>81,712</b>
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)							<b>257,000</b>
K. RESIDUAL FUNDS							<b>0</b>
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)							\$ <b>257,000</b> \$
M. COST SHARING PROPOSED LEVEL \$ <b>0</b>				AGREED LEVEL IF DIFFERENT \$			
PI/PI NAME <b>Michael J Khoo</b>				FOR NSF USE ONLY			
ORG. REP. NAME* <b>Margaret vigiolto</b>				INDIRECT COST RATE VERIFICATION			
		Date Checked	Date Of Rate Sheet			Initials - ORG	

# SUMMARY PROPOSAL BUDGET Cumulative

ORGANIZATION <b>Drexel University</b>				FOR NSF USE ONLY		
				PROPOSAL NO.	DURATION (months)	
PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR <b>Michael J Khoo</b>				AWARD NO.	Proposed	Granted
					NSF Funded Person-months	
A. SENIOR PERSONNEL: PI/PI, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)				CAL	ACAD	SUMR
1. <b>Michael J Khoo - PI</b>				0.00	0.00	2.00
2. <b>Eileen Abels - Co-PI</b>				0.00	0.00	0.50
3. <b>Sean Goggins - Co-PI</b>				0.00	0.00	3.00
4. <b>Jiexun Li - Co-PI</b>				0.00	0.00	2.00
5. <b>Xia Lin - Senior personnel</b>				0.00	0.00	0.00
6. ( <b>1</b> ) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)				0.00	0.00	0.50
7. ( <b>6</b> ) TOTAL SENIOR PERSONNEL (1 - 6)				0.00	0.00	8.00
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)						
1. ( <b>0</b> ) POST DOCTORAL SCHOLARS				0.00	0.00	0.00
2. ( <b>3</b> ) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)				12.00	0.00	0.00
3. ( <b>6</b> ) GRADUATE STUDENTS						
4. ( <b>0</b> ) UNDERGRADUATE STUDENTS						
5. ( <b>0</b> ) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)						
6. ( <b>0</b> ) OTHER						
TOTAL SALARIES AND WAGES (A + B)						
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)						
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)						
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING \$5,000.)						
TOTAL EQUIPMENT						
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)						
2. FOREIGN						
F. PARTICIPANT SUPPORT COSTS						
1. STIPENDS \$ _____ <b>0</b>						
2. TRAVEL _____ <b>0</b>						
3. SUBSISTENCE _____ <b>0</b>						
4. OTHER _____ <b>0</b>						
TOTAL NUMBER OF PARTICIPANTS ( <b>0</b> ) TOTAL PARTICIPANT COSTS						
G. OTHER DIRECT COSTS						
1. MATERIALS AND SUPPLIES						
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION						
3. CONSULTANT SERVICES						
4. COMPUTER SERVICES						
5. SUBAWARDS						
6. OTHER						
TOTAL OTHER DIRECT COSTS						
H. TOTAL DIRECT COSTS (A THROUGH G)						
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE)						
TOTAL INDIRECT COSTS (F&A)						
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)						
K. RESIDUAL FUNDS						
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)						
M. COST SHARING PROPOSED LEVEL \$ <b>0</b> AGREED LEVEL IF DIFFERENT \$						
PI/PI NAME <b>Michael J Khoo</b>				FOR NSF USE ONLY		
ORG. REP. NAME* <b>Margaret vigiolto</b>				INDIRECT COST RATE VERIFICATION		
				Date Checked	Date Of Rate Sheet	Initials - ORG

C \*ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET

## **Budget Justification**

### **Year 1**

\$10,263 is requested for 1 month summer support for Khoo; \$10,222 is requested for 1 month summer support for Goggins; and \$10,111 is requested for 1 month summer support for Li. Khoo, Goggins and Li will begin the work of developing the CECM tool. Khoo will work on the metadata administration tool, Li will work on the metadata recommender tool, and Goggins will work on the design of the user logs that will capture the interaction data. Khoo will also coordinate the project activities. \$6,517 is requested for 0.5 month summer support for Stahl; Stahl will provide initial liaison with the VMT project. \$6,032 is requested for 0.5 month summer support for Lin; Lin will help with the development of the metadata recommender tool, and with developing curriculum modules that support the tool's use in graduate classes.

\$19,800 is requested to support a programmer to adopt the VMT environment for the collaborative metadata creation tool. \$50,000 is requested to support to graduate student research assistants, who will support the PI and Co-PIs in the project development.

\$20,269 is requested for fringe benefits for relevant project staff. \$1,600 is requested for RA health insurance.

\$9,600 is requested for 2 domestic trips each for Khoo, Goggins and Li (=\$1,600 per trip), to travel to professional meetings to discuss and present the initial stages of the work.

\$77,261 is requested for 53.5% indirect costs on modified total direct costs of \$144,414.

### **Year 2**

\$10,777 is requested for 1 month summer support for Khoo; \$10,734 is requested for 1 month summer support for Goggins; and \$10,617 is requested for 1 month summer support for Li. Khoo, Goggins and Li will continue the work of developing the CECM tool. Khoo and Goggins will begin preliminary data analyses. Li will refine the algorithms and functionality of the metadata recommender tool. Khoo will also coordinate the project activities. \$7,229 is requested for 0.5 month summer support for Abels, who will oversee integration with the IPL and also develop curriculum modules that support the tool's use in graduate classes.

\$20,790 is requested to support a programmer to continue adopt and support the VMT environment for the collaborative metadata creation tool. \$52,500 is requested to support to graduate student research assistants, who will support the PI and Co-PIs in the project development.

\$19,368 is requested for fringe benefits for relevant project staff. \$1,680 is requested for RA health insurance.

\$10,080 is requested for 2 domestic trips each for Khoo, Goggins and Li (=\$1,680 per trip), to travel to professional meetings to discuss and present the initial stages of the work.

\$77,998 is requested for 54.25% indirect costs on modified total direct costs of \$143,775.

### **Year 3**

\$11,316 is requested for 1 month summer support for Khoo; \$11,271 is requested for 1 month summer support for Goggins; and \$11,148 is requested for 1 month summer support for Li. Khoo, Goggins and Li will continue the work of developing the CECM tool. Khoo and Goggins will continue the data analyses. Li will evaluate the success of the metadata recommender tool. Khoo, Goggins and Li will begin preparing research findings for presentation publication. Khoo will coordinate the project activities.

\$21,830 is requested to support a programmer to continue support the VMT environment for the collaborative metadata creation tool. \$55,125 is requested to support to graduate student research assistants, who will support the PI and Co-PIs in the project development.

\$17,892 is requested for fringe benefits for relevant project staff. \$1,764 is requested for RA health insurance.

\$10,584 is requested for 2 domestic trips each for Khoo, Goggins and Li (=\$1,764 per trip), and \$9,000 is requested for 1 international trip each for Khoo, Goggins and Li (=\$3,000 per trip), to travel to professional meetings to discuss and present the initial stages of the work.

\$81,712 is requested for 54.5% indirect costs on modified total direct costs of \$149,930.

## Current and Pending Support

(See GPG Section II.C.2.h for guidance on information to include on this form.)

The following information should be provided for each investigator and other senior personnel. Failure to provide this information may delay consideration of this proposal.	
Investigator: Michael Khoo	Other agencies (including NSF) to which this proposal has been/will be submitted.
<p>Support:    <input checked="" type="checkbox"/> Current    <input type="checkbox"/> Pending    <input type="checkbox"/> Submission Planned in Near Future    <input type="checkbox"/> *Transfer of Support</p> <p>Project/Proposal Title:    Addressing the 'metadata bottleneck' by developing and evaluating an online tool to support non-specialists to evaluate Dublin Core metadata records</p> <p>Source of Support:    Online Computer Library Center (OCLC)</p> <p>Total Award Amount: \$    14,855 Total Award Period Covered:    01/01/09 - 12/31/09</p> <p>Location of Project:    Drexel University</p> <p>Person-Months Per Year Committed to the Project.    Cal:0.00    Acad: 0.00    Sumr: 0.50</p>	
<p>Support:    <input type="checkbox"/> Current    <input checked="" type="checkbox"/> Pending    <input type="checkbox"/> Submission Planned in Near Future    <input type="checkbox"/> *Transfer of Support</p> <p>Project/Proposal Title:    Collaborative Knowledge Work in Socially Intelligent Computational Systems</p> <p>Source of Support:    NSF - SoCS</p> <p>Total Award Amount: \$    747,599 Total Award Period Covered:    01/01/10 - 12/31/12</p> <p>Location of Project:    Drexel University</p> <p>Person-Months Per Year Committed to the Project.    Cal:0.00    Acad: 0.00    Sumr: 1.00</p>	
<p>Support:    <input type="checkbox"/> Current    <input type="checkbox"/> Pending    <input type="checkbox"/> Submission Planned in Near Future    <input type="checkbox"/> *Transfer of Support</p> <p>Project/Proposal Title:</p> <p>Source of Support:</p> <p>Total Award Amount: \$    Total Award Period Covered:</p> <p>Location of Project:</p> <p>Person-Months Per Year Committed to the Project.    Cal:    Acad:    Sumr:</p>	
<p>Support:    <input type="checkbox"/> Current    <input type="checkbox"/> Pending    <input type="checkbox"/> Submission Planned in Near Future    <input type="checkbox"/> *Transfer of Support</p> <p>Project/Proposal Title:</p> <p>Source of Support:</p> <p>Total Award Amount: \$    Total Award Period Covered:</p> <p>Location of Project:</p> <p>Person-Months Per Year Committed to the Project.    Cal:    Acad:    Sumr:</p>	
<p>Support:    <input type="checkbox"/> Current    <input type="checkbox"/> Pending    <input type="checkbox"/> Submission Planned in Near Future    <input type="checkbox"/> *Transfer of Support</p> <p>Project/Proposal Title:</p> <p>Source of Support:</p> <p>Total Award Amount: \$    Total Award Period Covered:</p> <p>Location of Project:</p> <p>Person-Months Per Year Committed to the Project.    Cal:    Acad:    Summ:</p>	

\*If this project has previously been funded by another agency, please list and furnish information for immediately preceding funding period.

## Current and Pending Support

**(See GPG Section II.C.2.h for guidance on information to include on this form.)**

The following information should be provided for each investigator and other senior personnel. Failure to provide this information may delay consideration of this proposal.

<b>Investigator:</b> Eileen Abels	Other agencies (including NSF) to which this proposal has been/will be submitted.
-----------------------------------	---

<b>Support:</b> <input type="checkbox"/> Current <input checked="" type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support	
<b>Project/Proposal Title:</b> Expanding Access to Health Information Through Augmented Digital Reference Service	
<b>Source of Support:</b> IMLS (Federal)	
<b>Total Award Amount:</b> \$ 840,176 <b>Total Award Period Covered:</b> 10/01/09 - 09/30/12	
<b>Location of Project:</b> Drexel University	
<b>Person-Months Per Year Committed to the Project.</b> Cal:0.00 Acad: 0.00 Sumr: 1.00	

<b>Support:</b> <input type="checkbox"/> Current <input checked="" type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support	
<b>Project/Proposal Title:</b> Integrating Subject Categories of the Internet Public Library (IPL) and Librarians' Internet Index (LII)	
<b>Source of Support:</b> IMLS (Federal)	
<b>Total Award Amount:</b> \$ 305,360 <b>Total Award Period Covered:</b> 10/01/09 - 09/30/11	
<b>Location of Project:</b> Drexel University	
<b>Person-Months Per Year Committed to the Project.</b> Cal:0.00 Acad: 0.00 Sumr: 0.50	

<b>Support:</b> <input checked="" type="checkbox"/> Current <input type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support	
<b>Project/Proposal Title:</b> Transforming the IPL into a Virtual Learning Laboratory	
<b>Source of Support:</b> IMLS (Federal)	
<b>Total Award Amount:</b> \$ 613,513 <b>Total Award Period Covered:</b> 07/01/07 - 03/31/10	
<b>Location of Project:</b> Drexel University	
<b>Person-Months Per Year Committed to the Project.</b> Cal:0.00 Acad: 0.00 Sumr: 1.00	

<b>Support:</b> <input type="checkbox"/> Current <input checked="" type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support	
<b>Project/Proposal Title:</b> Collaborative Knowledge Work in Socially Intelligent Computational Systems	
<b>Source of Support:</b> NSF - SoCS	
<b>Total Award Amount:</b> \$ 747,599 <b>Total Award Period Covered:</b> 01/01/10 - 12/31/12	
<b>Location of Project:</b> Drexel University	
<b>Person-Months Per Year Committed to the Project.</b> Cal:0.00 Acad: 0.00 Sumr: 0.17	

<b>Support:</b> <input type="checkbox"/> Current <input type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support	
<b>Project/Proposal Title:</b>	
<b>Source of Support:</b>	
<b>Total Award Amount:</b> \$ Total Award Period Covered:	
<b>Location of Project:</b>	
<b>Person-Months Per Year Committed to the Project.</b> Cal: Acad: Summ:	

\*If this project has previously been funded by another agency, please list and furnish information for immediately preceding funding period.



## Current and Pending Support

(See GPG Section II.C.2.h for guidance on information to include on this form.)

The following information should be provided for each investigator and other senior personnel. Failure to provide this information may delay consideration of this proposal.	
Investigator: Sean Goggins	Other agencies (including NSF) to which this proposal has been/will be submitted.
Support: <input type="checkbox"/> Current <input checked="" type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support Project/Proposal Title: Collaborative Knowledge Work in Socially Intelligent Computational Systems  Source of Support:      NSF - SoCS Total Award Amount: \$    747,599 Total Award Period Covered:    01/01/10 - 12/31/12 Location of Project:      Drexel University Person-Months Per Year Committed to the Project.    Cal:0.00    Acad: 0.00    Sumr: 1.00	
Support: <input type="checkbox"/> Current <input type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support Project/Proposal Title:  Source of Support: Total Award Amount: \$                      Total Award Period Covered: Location of Project: Person-Months Per Year Committed to the Project.    Cal:            Acad:            Sumr:	
Support: <input type="checkbox"/> Current <input type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support Project/Proposal Title:  Source of Support: Total Award Amount: \$                      Total Award Period Covered: Location of Project: Person-Months Per Year Committed to the Project.    Cal:            Acad:            Sumr:	
Support: <input type="checkbox"/> Current <input type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support Project/Proposal Title:  Source of Support: Total Award Amount: \$                      Total Award Period Covered: Location of Project: Person-Months Per Year Committed to the Project.    Cal:            Acad:            Sumr:	
Support: <input type="checkbox"/> Current <input type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support Project/Proposal Title:  Source of Support: Total Award Amount: \$                      Total Award Period Covered: Location of Project: Person-Months Per Year Committed to the Project.    Cal:            Acad:            Sumr:	

\*If this project has previously been funded by another agency, please list and furnish information for immediately preceding funding period.

## Current and Pending Support

(See GPG Section II.C.2.h for guidance on information to include on this form.)

The following information should be provided for each investigator and other senior personnel. Failure to provide this information may delay consideration of this proposal.

Investigator: Jiexun Li

Other agencies (including NSF) to which this proposal has been/will be submitted.

Support:  Current  Pending  Submission Planned in Near Future  \*Transfer of Support  
Project/Proposal Title: A Framework of Automatic Hypothesis Generation for Clinical Research Agency

Source of Support: NIH  
Total Award Amount: \$ 798,423 Total Award Period Covered: 10/01/09 - 09/30/11  
Location of Project: Drexel University  
Person-Months Per Year Committed to the Project. Cal:1.00 Acad: 1.40 Sumr: 0.00

Support:  Current  Pending  Submission Planned in Near Future  \*Transfer of Support  
Project/Proposal Title: Collaborative Knowledge Work in Socially Intelligent Computational Systems

Source of Support: NSF - SoCS  
Total Award Amount: \$ 747,599 Total Award Period Covered: 01/01/10 - 12/31/12  
Location of Project: Drexel University  
Person-Months Per Year Committed to the Project. Cal:0.00 Acad: 0.00 Sumr: 1.00

Support:  Current  Pending  Submission Planned in Near Future  \*Transfer of Support  
Project/Proposal Title:

Source of Support:  
Total Award Amount: \$ Total Award Period Covered:  
Location of Project:  
Person-Months Per Year Committed to the Project. Cal: Acad: Sumr:

Support:  Current  Pending  Submission Planned in Near Future  \*Transfer of Support  
Project/Proposal Title:

Source of Support:  
Total Award Amount: \$ Total Award Period Covered:  
Location of Project:  
Person-Months Per Year Committed to the Project. Cal: Acad: Sumr:

Support:  Current  Pending  Submission Planned in Near Future  \*Transfer of Support  
Project/Proposal Title:

Source of Support:  
Total Award Amount: \$ Total Award Period Covered:  
Location of Project:  
Person-Months Per Year Committed to the Project. Cal: Acad: Sumr:

\*If this project has previously been funded by another agency, please list and furnish information for immediately preceding funding period.

## Current and Pending Support

(See GPG Section II.C.2.h for guidance on information to include on this form.)

The following information should be provided for each investigator and other senior personnel. Failure to provide this information may delay consideration of this proposal.

Investigator: Gerry Stahl	Other agencies (including NSF) to which this proposal has been/will be submitted.
<p>Support:   <input checked="" type="checkbox"/> Current   <input type="checkbox"/> Pending   <input type="checkbox"/> Submission Planned in Near Future   <input type="checkbox"/> *Transfer of Support</p> <p>Project/Proposal Title:   Dynamic Support for Virtual Math Teams</p> <p>Source of Support:       National Science Foundation</p> <p>Total Award Amount: \$    306,355 Total Award Period Covered:    08/01/09 - 07/31/12</p> <p>Location of Project:     Drexel University</p> <p>Person-Months Per Year Committed to the Project.   Cal:0.00   Acad: 0.00   Sumr: 1.50</p>	
<p>Support:   <input type="checkbox"/> Current   <input checked="" type="checkbox"/> Pending   <input type="checkbox"/> Submission Planned in Near Future   <input type="checkbox"/> *Transfer of Support</p> <p>Project/Proposal Title:   Collaborative Knowledge Work in Socially Intelligent Computational Systems</p> <p>Source of Support:       NSF - SoCS</p> <p>Total Award Amount: \$    747,599 Total Award Period Covered:    01/01/10 - 12/31/12</p> <p>Location of Project:     Drexel University</p> <p>Person-Months Per Year Committed to the Project.   Cal:0.00   Acad: 0.00   Sumr: 0.17</p>	
<p>Support:   <input type="checkbox"/> Current   <input type="checkbox"/> Pending   <input type="checkbox"/> Submission Planned in Near Future   <input type="checkbox"/> *Transfer of Support</p> <p>Project/Proposal Title:</p> <p>Source of Support:</p> <p>Total Award Amount: \$                    Total Award Period Covered:</p> <p>Location of Project:</p> <p>Person-Months Per Year Committed to the Project.   Cal:           Acad:           Sumr:</p>	
<p>Support:   <input type="checkbox"/> Current   <input type="checkbox"/> Pending   <input type="checkbox"/> Submission Planned in Near Future   <input type="checkbox"/> *Transfer of Support</p> <p>Project/Proposal Title:</p> <p>Source of Support:</p> <p>Total Award Amount: \$                    Total Award Period Covered:</p> <p>Location of Project:</p> <p>Person-Months Per Year Committed to the Project.   Cal:           Acad:           Sumr:</p>	
<p>Support:   <input type="checkbox"/> Current   <input type="checkbox"/> Pending   <input type="checkbox"/> Submission Planned in Near Future   <input type="checkbox"/> *Transfer of Support</p> <p>Project/Proposal Title:</p> <p>Source of Support:</p> <p>Total Award Amount: \$                    Total Award Period Covered:</p> <p>Location of Project:</p> <p>Person-Months Per Year Committed to the Project.   Cal:           Acad:           Summ:</p>	

\*If this project has previously been funded by another agency, please list and furnish information for immediately preceding funding period.

## Current and Pending Support

(See GPG Section II.C.2.h for guidance on information to include on this form.)

The following information should be provided for each investigator and other senior personnel. Failure to provide this information may delay consideration of this proposal.	
Investigator: Xia Lin	Other agencies (including NSF) to which this proposal has been/will be submitted.
Support: <input checked="" type="checkbox"/> Current <input type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support Project/Proposal Title: Developing Faculty in Digital Librarianship for the 21st Century  Source of Support: IMLS (Federal) Total Award Amount: \$ 992,100 Total Award Period Covered: 09/01/08 - 08/31/12 Location of Project: Drexel University Person-Months Per Year Committed to the Project. Cal:0.00 Acad: 1.00 Sumr: 0.30	
Support: <input checked="" type="checkbox"/> Current <input type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support Project/Proposal Title: Planning Grant: I/UCRC Center for Visual Decision Informatics  Source of Support: NSF (Federal) Total Award Amount: \$ 10,000 Total Award Period Covered: 09/01/09 - 08/31/10 Location of Project: Drexel University Person-Months Per Year Committed to the Project. Cal:0.00 Acad: 0.00 Sumr: 0.00	
Support: <input checked="" type="checkbox"/> Current <input type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support Project/Proposal Title: Interactive Visual Query Refinement and Recommendation based on automatic topic detection  Source of Support: NSF (Federal) - NSDL Program Total Award Amount: \$ 0 Total Award Period Covered: 01/01/00 - 01/01/00 Location of Project: Drexel University Person-Months Per Year Committed to the Project. Cal:0.50 Acad: 0.00 Sumr: 0.00	
Support: <input checked="" type="checkbox"/> Current <input type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support Project/Proposal Title: A Framework of Automatic Hypotheses Generation for Emergency Medicine Research  Source of Support: NIH (Federal) Total Award Amount: \$ 0 Total Award Period Covered: 01/01/00 - 01/01/00 Location of Project: Person-Months Per Year Committed to the Project. Cal:0.50 Acad: 0.00 Sumr: 0.00	
Support: <input type="checkbox"/> Current <input checked="" type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support Project/Proposal Title: Collaborative Knowledge Work in Socially Intelligent Computational Systems  Source of Support: NSF - SoCS Total Award Amount: \$ 747,599 Total Award Period Covered: 01/01/10 - 12/31/12 Location of Project: Drexel University Person-Months Per Year Committed to the Project. Cal:0.00 Acad: 0.00 Summ: 0.17	
*If this project has previously been funded by another agency, please list and furnish information for immediately preceding funding period.	

## Current and Pending Support

(See GPG Section II.C.2.h for guidance on information to include on this form.)

The following information should be provided for each investigator and other senior personnel. Failure to provide this information may delay consideration of this proposal.	
Investigator: <b>Baba Kofi Weusijana</b>	Other agencies (including NSF) to which this proposal has been/will be submitted.
Support: <input type="checkbox"/> Current <input checked="" type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support Project/Proposal Title: <b>Collaborative Knowledge Work in Socially Intelligent Computational Systems</b>	
Source of Support: <b>NSF - SoCS</b> Total Award Amount: \$ <b>747,599</b> Total Award Period Covered: <b>01/01/10 - 12/31/12</b> Location of Project: <b>Drexel University</b> Person-Months Per Year Committed to the Project. <b>Cal:4.00    Acad: 0.00    Sumr: 0.00</b>	
Support: <input type="checkbox"/> Current <input type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support Project/Proposal Title:	
Source of Support: Total Award Amount: \$                      Total Award Period Covered: Location of Project: Person-Months Per Year Committed to the Project.    Cal:              Acad:              Sumr:	
Support: <input type="checkbox"/> Current <input type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support Project/Proposal Title:	
Source of Support: Total Award Amount: \$                      Total Award Period Covered: Location of Project: Person-Months Per Year Committed to the Project.    Cal:              Acad:              Sumr:	
Support: <input type="checkbox"/> Current <input type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support Project/Proposal Title:	
Source of Support: Total Award Amount: \$                      Total Award Period Covered: Location of Project: Person-Months Per Year Committed to the Project.    Cal:              Acad:              Sumr:	
Support: <input type="checkbox"/> Current <input type="checkbox"/> Pending <input type="checkbox"/> Submission Planned in Near Future <input type="checkbox"/> *Transfer of Support Project/Proposal Title:	
Source of Support: Total Award Amount: \$                      Total Award Period Covered: Location of Project: Person-Months Per Year Committed to the Project.    Cal:              Acad:              Summ:	

\*If this project has previously been funded by another agency, please list and furnish information for immediately preceding funding period.

## FACILITIES, EQUIPMENT & OTHER RESOURCES

---

**FACILITIES:** Identify the facilities to be used at each performance site listed and, as appropriate, indicate their capacities, pertinent capabilities, relative proximity, and extent of availability to the project. Use "Other" to describe the facilities at any other performance sites listed and at sites for field studies. USE additional pages as necessary.

**Laboratory:** We will have full access to an HCI user-testing lab at the iSchool, Drexel University. The lab can be used for intensive, fine-grained study of users' interactions with the metadata tool.

**Clinical:**

**Animal:**

**Computer:** We will have full access to all the necessary computer equipment needed to complete this project at the iSchool at Drexel. This includes (but is not limited to) computers; servers; data storage and backup, and other facilities; and necessary software.

**Office:** We will have full access to all necessary office infrastructure needed to complete the project at Drexel University. This includes (but is not limited to) personal offices; conference rooms; and telephone, fax, and copiers.

**Other:**

---

**MAJOR EQUIPMENT:** List the most important items available for this project and, as appropriate identifying the location and pertinent capabilities of each.

---

**OTHER RESOURCES:** Provide any information describing the other resources available for the project. Identify support services such as consultant, secretarial, machine shop, and electronics shop, and the extent to which they will be available for the project. Include an explanation of any consortium/contractual arrangements with other organizations.

We will have full access to all necessary office support staff needed to complete the project at Drexel University. This includes (but is not limited to) secretarial assistance, technical and computer support, and other logistical staff.

---