

# IDEA Institute on Artificial Intelligence Program

July 10 - July 15, 2022

The University of Texas at Austin  
School of Information



**Dania Bilal, Clara M. Chu, Soo Young Rieh, Nicole Coleman, Claudia Engel, Jiangen He, William H. Mischo**



Dania Bilal, Clara M. Chu, Soo Young Rieh, Nicole Coleman,  
Claudia Engel, Jiangen He, William H. Mischo, 2022

## Home Institutions of the IDEA Institute on AI



## Partnered with



## Funded by



This project was made possible in part by the [Institute of Museum and Library Services](#) grant: [RE-246419-OLS-20](#)

© 2022 by Dania Bilal, Clara M. Chu, Soo Young Rieh, Nicole Coleman, Claudia Engel, Jiangen He, William H. Mischo. This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0 Deed). To view a copy of this license, visit: <https://creativecommons.org/licenses/by-nc-sa/4.0/>

### *Suggested citation:*

Bilal, D., Chu, C.M., Rieh, S.Y., Coleman, N., Engel, C., He, J., & Mischo, W.H. (2022). *IDEA Institute on Artificial Intelligence Program*. <https://idea.infosci.utk.edu/pubs-talks/>

IDEA Institute on Artificial Intelligence

<https://idea.infosci.utk.edu/>

# Table of Contents

<b>Schedule at a Glance</b>	<b>3</b>
<b>Groups</b>	<b>4</b>
<b>Sunday, July 10, 2022</b>	<b>5</b>
Program Introduction	5
Survey of Existing LAM Projects and Efforts	6
Handout - Overview of Existing AI Projects	7
AI in Information Search and Discovery	9
Ethical Considerations and Guidelines	11
Small Group Activity: Fellows' AI Project Idea Exchange	12
<b>Monday, July 11, 2022</b>	<b>13</b>
What Machine Learning Can Do	13
Project Planning I: Project Design	14
Machine Learning and Coding	15
Invited Talk	16
<b>Tuesday, July 12, 2022</b>	<b>17</b>
Project Planning II: Data Collection, Classification, and Transformation	17
Project Planning III: Roles & Implementation	19
Reflection on Fellows' Learning and Institute Experience	20
Harvesting, Evaluating, and Training Data Sets for Use in AI	21
Equity, Diversity, Inclusion and Accessibility in Fellows' AI projects	22
<b>Wednesday, July 13, 2022</b>	<b>26</b>
Conversational AI Part 1 - Theoretical	26
Linked Open Data (LOD)	27
Machine Learning for Text with Topic Modeling and Clustering	28
Networking Dinner with Austin AI Experts	29
<b>Thursday, July 14, 2022</b>	<b>30</b>
Conversational AI Part 2 - Application	30
Post-Program Feedback and Activities	31
<b>Friday, July 15, 2022</b>	<b>33</b>
2022 Fellows AI Project and Action Plan	33
Capstone Showcase Guidelines	35
Capstone: Appendix A: SMARTIE Goals Worksheet	36
Capstone: Appendix B: Developing an Action Plan	38
<b>Appendix: 2022 Institute Information Sheet</b>	<b>39</b>

Sunday July 10	Monday July 11	Tuesday July 12	Wednesday July 13	Thursday July 14	Friday July 15
<p>9 – 10 AM Welcome by Eric T. Meyer, Dean, School of Information   Introductions &amp; Program overview Dania Bilal 10:00 – 11:00 AM Survey of existing LAM projects and efforts Nicole Coleman &amp; Claudia Engel</p>	<p>8:30 – 10:00 AM What machine learning can do Nicole Coleman &amp; Claudia Engel</p>	<p>8:30 – 10:30 AM AI Project Planning, Part 2 Data collection, classification, and transformation Nicole Coleman &amp; Claudia Engel</p>	<p>8:30 – 10:30 AM Conversational AI, Part 1 Theoretical Jianguo He</p>	<p>8:30 – 11:45 AM Conversation AI, Part 2 Applications Jianguo He</p>	<p>9 AM – 12 PM Capstone Showcase (open to local library and information professionals and supervisors) Certificates and Farewell</p>
<p>11:15 AM – 12:15 PM AI in information search and discovery Bill Mischo</p>	<p>10:15 AM – 12:00 PM AI project planning, Part 1 – project design N.Coleman &amp; C. Engel</p>	<p>10:45 – 11:45 AM AI Project planning, Part 3: Roles and implementation N. Coleman &amp; C. Engel</p>	<p>10:45 – 12:00 AM Linked Open Data Bill Mischo</p>	<p>11:45 AM – 12:00 PM Post-Program Questionnaire and Activities</p>	
<p>12:15 – 1:15 PM Lunch</p>	<p>12:00– 1:00 PM Lunch</p>	<p>11:45 – 1:00 PM Reflection on Fellows’ learning experience Soo Young Rieh</p>	<p>12:00 – 1:00 PM Lunch</p>	<p>12:00 – 1:00 PM Lunch</p>	<p>Travel home</p>
<p>1:15 – 3:30PM Ethical considerations and guidelines Nicole Coleman &amp; Claudia Engel</p>	<p>1:00 – 4:30 PM Machine Learning and coding Jianguo He</p>	<p>1:00 – 3:00 PM Harvesting, evaluating and training data sets for use in AI Jianguo He</p>	<p>1:00 – 3:00 PM Machine learning for text with topic modeling and clustering Bill Mischo</p>	<p>1:00 – 4:00 PM Fellows work on their own project</p>	
<p>3:45 – 5:00 PM Small Group Activity: Fellows’ AI Project Idea Exchange</p>		<p>3:15 – 4:30 PM Equity, Diversity, Inclusion and Accessibility in Fellows’ AI projects Clara M. Chu</p>	<p>3:15 – 4:45 PM Fellows work on their own project</p>	<p>4:00 – 5:15 PM UT Perry-Castañeda Library Tour Michael Shensky</p>	
	<p>5:30 – 6:30 PM Invited Speaker: Jason Griffey via Zoom 6:30 - 8:00 PM Networking Dinner @ Hilton Garden Inn 2nd Floor Intro Room</p>	<p>4:30 – 6:00 PM Coded Bias (optional film viewing)</p>	<p>5:30 – 7:30 PM Expert Panel and Networking Dinner @ Hilton Garden Inn 17th Floor Apex Rooftop</p>		

## Groups

Group	Fellows	Position	Affiliation
AI Ethics (Claudia Engel, Clara Chu)	Sandy Avila	Science Librarian	University of Central Florida
	Kimberly Grotewold	Education Librarian	Texas A&M University-San Antonio
	Valerie Lookingbill	Sciences Librarian	University of South Carolina
	Lencia Beltran	Open Science Coordinator	Carnegie Mellon University
	Sharon Whitfield	Electronic Resource & User Access Librarian	Rider University
User services (Nicole Coleman, Yujin Choi)	Kineret Ben-Knaan	Research and Assessment Librarian	University of Miami
	Derek Braeden	Web Developer	Salt Lake City Public Library
	Sharon Clapp	Digital Resources Librarian	CCSU Library
	Robert Wilson	Systems and Analytics Librarian	Middle Tennessee State University
	Borui Zhang	Natural language processing specialist	University of Florida
Collection use (Bill Mischo, Dania Bilal)	Thomas Ferrill	Head of Creative Spaces	University of Utah, Marriott Library
	Mary Holm	Cataloging Metadata Librarian	Virginia State University
	Rafia Mirza	Humanities Librarian (Digital Humanities)	Southern Methodist University
	Martha Roseberry	Applications Analyst	Virginia Commonwealth University
	Daniella Smith	Professor	University of North Texas
Data Management & ML techniques (Jianguan He, Soo Young Rieh)	Mohammad AlHamad	E-Resource Strategist	Missouri State University
	Mary Aycock	Database and Metadata Management Librarian	Texas State University
	Erik Radio	Metadata Librarian	University of Colorado Boulder
	Michael Shensky	GIS & Geospatial Data Coordinator	University of Texas at Austin
	Raymond Wang	Reference/Instruction Coordinator	Community College of Baltimore County



## Program Introduction and Agreements

### Purpose (or learning goals):

- Welcome (Soo Young Rieh)
- Welcome Remarks by Dr. Eric T. Meyer, Dean, School of Information
- Institute overview (Dania Bilal)

**Date and Time:** Sunday July 10, 9:00 AM – 10:00 AM

**Materials:** Institute Overview, Program, and Information Sheet

**Who:** Soo Young Rieh and Dania Bilal

**How:** Presentation and Q&A

### Resources:

Institute website: <https://idea.infosci.utk.edu/institute2021/institute2022/> | Box folder for Institute materials: [go.illinois.edu/ideainstitute2022](https://go.illinois.edu/ideainstitute2022)

### PROGRAM AGREEMENTS

#### Use of Institute Materials:

2022 IDEA Institute on AI materials may be used by Institute Fellows and the Advisory Board for your learning and research purposes with the appropriate attribution. All other uses require permission by the Institute leadership team.

#### Talent Release:

All Institute participants are requested to sign the Talent Release form.



## Survey of existing LAM projects and efforts

**Purpose (or Learning Goals):** Introduction to projects and programs currently or previously developed by libraries, archives, and museums.

**Date and/or Time:** Sunday July 10, 10:00 AM – 11:00 AM

**Materials:** Slides

**Who:** Claudia Engel and Nicole Coleman

**How:** Discussion, working groups, and presentation.

- Which LAM organizations are currently engaged in AI projects?
- What problems do the organizations and projects try to solve?
- What are methods and workflows applied?
- What are the obstacles, shortcomings, limitations?
- What is the intent, scope, and sustainability plan for the project?

### References:

Collections as Data projects, [Jim Crow and Algorithms of Resistance](#)

Library of Congress Projects, [Newspaper Navigator](#)

Yale Beinecke Rare Book & Manuscript Library, [PixPlot](#)

Annif - [Tool for automated subject indexing and classification](#)

### Recommended Readings:

Wellcome Collection, [Image Pathways](#)

National Library of New South Wales, [Aereo](#)  
[Mauricio Giraldo presents on the making of Aereo](#)

HAMLET by Andromeda Yelton, [How About Machine Learning Enhancing Theses?](#)



**IDEA Institute on AI | July 10 – July 15, 2022  
Day One - Overview of Existing AI Projects**

Each group reviews one of the following projects:

1. Jim Crow and Algorithms of Resistance  
<https://onthebooks.lib.unc.edu/>
2. Newspaper Navigator  
<https://labs.loc.gov/work/experiments/newspaper-navigator/>
3. National Library of Norway Maken  
<https://www.nb.no/maken/>
4. Harvard Art Museums  
<https://ai.harvardartmuseums.org/>

And discuss the following questions:

**WHO** is engaged in the project?

**WHAT** is the problem the project tries to solve?

**HOW** does the data shape the project?





**WHAT** are the outcomes?

**CHALLENGES**

Are there obstacles, shortcomings, limitations and if so which ones?

**CONTEXT**

What is the intent, scope, and sustainability plan for the project?



## AI in Information Search and Discovery

### **Purpose (or learning goals):**

Discussion and demonstration of several AI and ML technologies being employed in information discovery and delivery systems and bibliometric systems

**Date and Time:** Sunday July 10, 11:15 AM – 12:15 PM

**Materials:** PowerPoint Slides, System Demonstrations, Exercise

**Who:** William H. Mischo, University of Illinois at Urbana-Champaign

**How:** Presentation, Demonstration, and Discussion

- Discovery at an inflection point
- Exciting time; we have the tools and data for custom solutions
- APIs are ubiquitous
- Proliferation of repositories: Dimensions, Lens, etc
- Semantic search implementations: YewNo, EBSCO's
- Illinois (and 45 other) Bento style discovery systems
- Recommender and automatic classification case studies
- User behaviors paramount
- AI and ML applications

Exercise

- Examine the Subject Recommendation Algorithm

### **Reference:** \_

University of Illinois Easy Search Bento implementation: <https://www.library.illinois.edu/>

OhioLINK and Ithaka vision for a new library system:

[It's not What Libraries Hold; It's Who Libraries Serve](#)

### **Recommended Readings:**

[Investigating the National Need for Library Based Topic Modeling Discovery Systems White Paper](#)

[Machine Learning, Libraries, and Cross-Disciplinary Research: Possibilities and Provocations](#)



## Ethical Considerations and Guidelines

**Purpose (or Learning Goals):** Review of ethical implications of and around AI

**Date and/or Time:** Sunday July 10, 1:15 PM – 3:30 PM

**Materials:** Slides

**Who:** Claudia Engel and Nicole Coleman

**How:** Discussion, working groups, and presentation.

1. What are your concerns around ethics and AI?
2. Topics in AI and Ethics (slides)
3. How are ethical concerns relevant for the libraries and your project in particular?
4. Exercise: Towards AI guiding values for LAM organizations (group work)

### References:

Ananny, Mike, and Kate Crawford. "Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability." *new media & society* 20.3 (2018): 973-989.

Benjamin, Ruha. "Race after technology: Abolitionist tools for the new jim code." *Social Forces* (2019).

Crawford, Kate. *The Atlas of AI*. Yale University Press, 2021.

Gray, Mary L., and Siddharth Suri. *Ghost work: How to stop Silicon Valley from building a new global underclass*. Eamon Dolan Books, 2019.

Kearns, Michael, and Aaron Roth. *The ethical algorithm: The science of socially aware algorithm design*. Oxford University Press, 2019.

Z. Epstein, B. H. Payne, J. H. Shen, A. Dubey, B. Felbo, M. Groh, N. Obradovich, M. Cebrian, I. Rahwan (2018). Closing the AI Knowledge Gap. arXiv:1803.07233 [cs.CY]

## Recommended Readings:

Boulamwini, Joy, star. Shalini Kantayya, director. [Coded Bias](#). 7th Empire media. November 11, 2020.

Benjamin, Ruha. "[Assessing risk, automating racism.](#)" *Science* 366.6464 (2019): 421-422.

Long, Hoyt. "[Can Algorithmic Bias Teach Us About Hate?](#)" Public Books. Online. June 28, 2021.

Noble, Safiya Umoja. *Algorithms of oppression*. New York University Press, 2018.

See also: [Algorithms of Oppression: How Search Engines Reinforce Racism - Dr. Safiya Noble](#)



## **Small Group Activity: Fellows' AI Project Idea Exchange**

**Purpose (or learning goals):**

Fellows will share their project ideas in a concise and efficient manner. After this session, each fellow will be able to continue conversations about common interests.

**Date and Time:** Sunday July 10, 3:45 PM – 5:00 PM

**Who:** Soo Young Rieh, Clara Chu, and Dania Bilal

**How:** Each fellow will try to answer the following questions.

1. What problem/need are you addressing?
2. What contributions will your project make to your organization?
3. What data do you need and is it available at your organization?
4. What AI techniques or technology do you think you'll need to use?
5. What skills and knowledge do you need to learn during the Institute?



## **What Machine Learning Can Do**

**Purpose (or Learning Goals):** Machine Learning in practice: Interactive experiments with AI

**Date and/or Time:** Monday July 11, 8:30 AM – 10:00 AM

**Materials:** Slides

**Who:** Claudia Engel and Nicole Coleman

**How:** Participants will interact with several online ML tools to train a model without coding.

**References:** This is supposed to be a fun experience and we prefer to not provide any pointers ahead of time.



## **Project Planning I: Project Design**

**Purpose (or Learning Goals):** Draft a design outline for an AI project

**Date and/or Time:** Monday July 11, 10:15 AM – 12:00 PM

**Materials:** Slides

**Who:** Claudia Engel and Nicole Coleman

**How:** Draft a Project Design Document

- Problem
  - What is your goal? Break the problem down into discrete tasks
  - Critical analysis of how to achieve the outcome you want.
  - What is the scope of the project? What is in, what is out?
  
- Exemplars
  - Find resources that describe applications of the techniques you plan to use.
  - Are there existing models you can use to train to your data?
  
- Methods
  - Often there may be several steps to accomplish your goal. What techniques are appropriate? How do you know?

**References:**

Methods planner



## Machine Learning and Coding

### Purpose:

This session provides not only a pipeline for understanding machine learning, but also python coding skills to implement machine learning skills. Topics for this session include: introduction to machine learning, kNN, linear regression model, logistic regression model, and other models.

Fellows will gain an understanding of machine learning and its implementation. Specific objectives are:

- Understand the capability of machine learning models.
- Design and implement the pipeline of a machine learning application.
- Discuss the role of data in a machine learning application.
- Apply data and fit machine learning models to the data.
- Select Python packages for implementing machine learning models.

**Date and/or Time:** Monday July 11, 1:00 PM – 4:30 PM

**Materials:** Slides, a Colab notebook for code examples, and a Colab notebook for exercise.

**Who:** Jiangen He, [jiangen@utk.edu](mailto:jiangen@utk.edu)

**How:** This session will use lectures and a hands-on exercise.

### References:

scikit-learn Tutorials <https://scikit-learn.org/stable/tutorial/index.html>





## Invited Talk

### Talk Title:

Machine Learning and the Research Ecosystem

**Talk:** Monday July 11, 5:30 – 6:30 PM

Networking Dinner @ Hilton Garden Inn, 2nd Floor Intro Room 6:40 - 8 PM

### Bio

[Jason Griffey](#) is the Director of Strategic Initiatives at NISO, where he works to identify new areas of the information ecosystem where standards expertise is useful and needed and leads ongoing projects such as NISO's participation in the Coalition for Seamless Access. He is also the Chair and Director of the NISO Plus conference (<http://niso.plus>). Prior to joining NISO in 2019, Jason ran his own technology consulting company for libraries, has been both an Affiliate at metaLAB and a [Fellow and Affiliate at the Berkman Klein Center for Internet & Society](#) at Harvard University, and was an academic librarian in roles ranging from reference and instruction to Head of Library IT at the University of TN at Chattanooga.

Jason has written extensively on technology and libraries, including multiple books and a series of full-periodical issues on technology topics, most recently [AI & Machine Learning in Libraries](#) and [Library Spaces and Smart Buildings: Technology, Metrics, and Iterative Design](#) from 2018. He has spoken internationally on topics such as artificial intelligence & machine learning, the future of technology and libraries, decentralization and the Blockchain, privacy, copyright, and intellectual property. A full list of his publications and presentations [can be found on his CV](#).

He can be stalked obsessively online (@griffey just about everywhere), and spends his free time with his daughter Eliza, reading, obsessing over gadgets, and preparing for the inevitable zombie uprising.



## **Project Planning II: Data collection, classification, and transformation**

**Purpose (or Learning Goals):** Discuss the role and importance of data in Machine learning.

**Date and/or Time:** Tuesday July 12, 8:30 AM – 10:30 AM

### **Materials: Slides**

**Who:** Claudia Engel and Nicole Coleman

### **How:**

- What is data?
  - Data formats: “raw,” unstructured, semi-structured, structured
  - Data curation
- Constructing a data biography
  - Data sources, provenance and statistics
- Transformations and classification
  - Metadata and labels, crowdsourcing
  - The interpretive work of annotation
- Data “bias” and feminist approaches

### **References:**

[The Point of Collection](#), Mimi Onuoha

[Data Biographies](#) , Heather Krause

### **Recommended Readings:**

D'ignazio, Catherine, and Lauren F. Klein. *Data feminism*. MIT press, 2020.

C. Bowker and Susan Leigh Star *Sorting Things Out: Classification and Its Consequences*, 1999

Jo, Eun Seo, and Timnit Gebru. "Lessons from archives: Strategies for collecting sociocultural data in machine learning." *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. 2020.

Loukissas, Yanni Alexander. *All data are local: Thinking critically in a data-driven society*. MIT Press, 2019.

Milagros Miceli, Martin Schuessler, Tianling Yang *Between Subjectivity and Imposition: Power Dynamics in Data Annotation for Computer Vision* (2020)

Paullada, Amandalynne, et al. "Data and its (dis) contents: A survey of dataset development and use in machine learning research." *arXiv preprint arXiv:2012.05345* (2020).



## **Project Planning III: Roles & Implementation**

**Purpose (or Learning Goals):** Who is involved in an AI project and what are their roles

**Date and/or Time:** Tuesday July 12, 10:45 AM – 11:45 AM

**Materials:** Slides

**Who:** Claudia Engel and Nicole Coleman

**How:**

Role Play Exercise:

- Who is on the team? What are the tasks?
- How do you make sure your audience is being served? Who is it serving? Who is it **not** serving?
- How do you communicate? What are your collaboration practices?



## Reflection on Fellows' Learning and Institute Experience

**Purpose (or learning goals):**

To understand Fellows' learning experiences at the Institute and identify the areas for improvement

**Date and Time:** Tuesday July 12, 11:45 AM – 1:00 PM

**Who:** Three Groups of Fellows; Each group discussion will be facilitated by Soo Young Rieh, Clara Chu, and Dania Bilal

**How:** Fellows will discuss the following topics.

What are the top two things that are going well at the Institute?

Please share an experience that you found particularly valuable during the onboarding or the Institute

Think back to your initial expectations of the Institute. Are your experiences matching these expectations? Why or why not?

Have you experienced any significant barriers to participating in the Institute so far?

Is there any feedback you would like to share about the Institute that we can improve now or for the future?



## Harvesting, evaluating, and training data sets for use in AI

### **Purpose:**

This session will discuss the methods for harvesting data, preparing training datasets, and evaluating AI models.

Fellows will have gained an understanding of data collection, data cleaning, training models with data, and various evaluation methods. Specific topics are:

- Open data sources
- Crawler scraping (social media and other web applications)
- Log collection (front-end capture and back-end script)
- Cleaning and validating data
- Preparing training datasets and training
- Metrics for evaluating trained model

**Date and/or Time:** Tuesday July 12, 1:00 PM – 3:00 PM

**Materials:** Slides, a Colab notebook for code examples, and a Colab notebook for exercise.

**Who:** Jiangen He, [jjangen@utk.edu](mailto:jjangen@utk.edu)

**How:** This session will use lectures, discussion, and a hands-on exercise.

**References:** Four Basic Ways to Automate Data Extraction

<https://towardsdatascience.com/four-basic-ways-to-automate-data-extraction-3151064dc110>



## Equity, Diversity, Inclusion and Accessibility in Fellows' AI projects

**Purpose:** Examine EDIA in the development (process and content) of Fellows' AI projects in their workplace.

*The IDEA Institute values equity, diversity and inclusion:*

The IDEA Institute on AI is committed to equity, diversity and inclusion (EDI) in all aspects of the program in order to deepen the knowledge and skills of library and information professionals in using artificial intelligence (AI) to enhance information access, discovery and services for users, independent of race, ethnicity, gender identity, sexual orientation, socio-economic status, age, physical abilities, and other dimensions of potential inequity.

Learn 4 aspects of AI:

- Conceptual (using UX AI human-centered interaction framework to guide the design of AI projects)
- Technical (introducing voice-enabled digital technology, ML algorithms, preparing data for ML, etc.)
- Social (covering ethical and AI impact, including algorithmic bias; privacy, safety, challenges, etc.)
- Applied (designing a variety of hands-on activities in AI and ML in collaboration with the Institute's instructors.)

**Date and Time:** Tuesday July 12, 3:15 PM – 4:30 PM

**Materials:** Project and Action Plan (available at [go.illinois.edu/ideainstitute2022](https://go.illinois.edu/ideainstitute2022))

**Who:**

Facilitator: Clara M. Chu, Director and Mortenson Distinguished Professor, Mortenson Center for International Library Programs, University of Illinois at Urbana-Champaign. Email: [cmchu@illinois.edu](mailto:cmchu@illinois.edu)

**How:**

**1. Assume EDIA work is mindful and intentional.**

*Our work doesn't just become equitable, diverse, inclusive and accessible; instead, we have to be intentional and plan for it.* (Review above: IDEA Institute on AI values and aspects of AI)

Diversity informs and guides the way we think, act and know as individuals or collectively as communities and organizations. Thus, in order to develop our cultural competence, we need to consider our Affective, Behavioral and Cognitive (ABC) development. (informed by *ALISE Diversity Statement*, 2013 <https://www.alise.org/alise---alise-diversity-statement>)

THINK (Affective)	ACT (Behavioral)	KNOW (Cognitive)
<p><i>Awareness of and analyzing one's internal biases, perspectives and belief system regarding cultural differences, and willingness to learn about cultures and systemic barriers.</i></p> <ul style="list-style-type: none"> <li>• accept diversity</li> <li>• respect difference</li> <li>• recognize commonalities</li> <li>• question cultural assumptions</li> <li>• be open-minded</li> <li>• acknowledge culture shapes worldview</li> <li>• accept change</li> <li>• foster inclusivity</li> </ul>	<p><i>Action and skills that promote diversity and inclusion, respect of cultural differences and social justice.</i></p> <ul style="list-style-type: none"> <li>• Make diversity and inclusion a priority</li> <li>• Invest in diversity</li> <li>• Shift the diversity paradigm from quantity to quality</li> <li>• Audit diversity in institution</li> <li>• Implement mechanisms to recognize and address biases and inequities</li> <li>• Measure institutional impact of diversity, e.g., Diversity Return on Investment</li> <li>• Incorporate diversity in the hiring, retention and promotion of faculty and staff</li> <li>• Recruit, retain and promote underrepresented students</li> <li>• Exhibit cross-cultural leadership</li> </ul>	<p>Knowledge of cultures, cultural differences and systemic barriers, and ways to reduce injustice.</p> <ul style="list-style-type: none"> <li>• cross-cultural practices and skills</li> <li>• Language diversity and its cultural implications</li> <li>• Inequities stem from structural differences and power</li> </ul>

2. Become familiar with concepts and terms. (see also References)

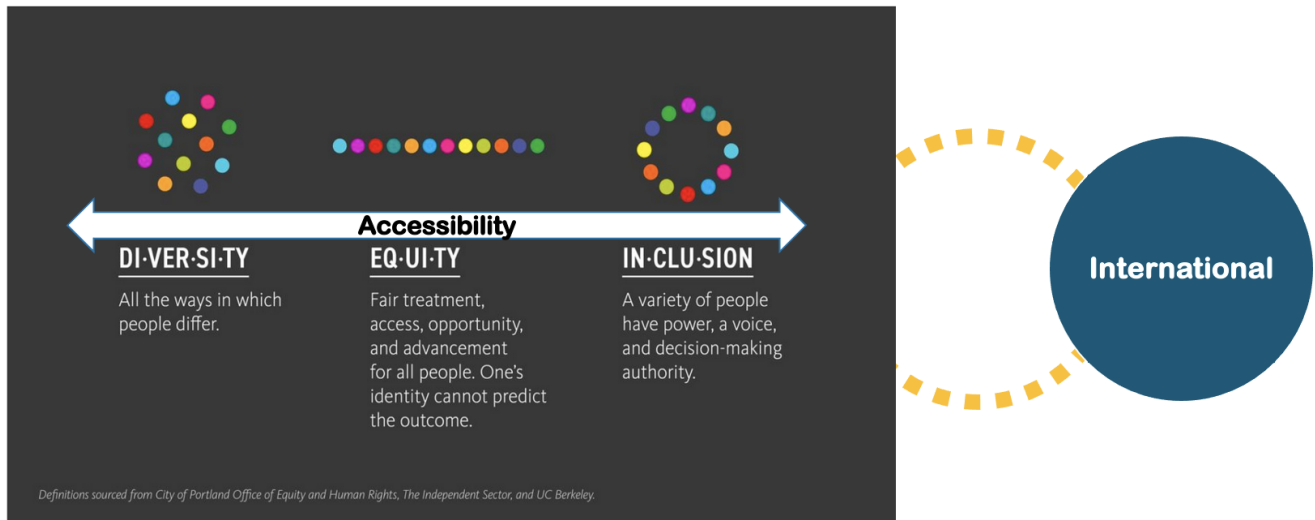


Image adapted from: <https://ideal.com/diversity-equity-inclusion/>



Three concepts of diversity offered by Sempra Energy (<http://www.sempra.com/diversity/dAbout.htm>), which distinguish individual, group, and institutional differences:

**Human diversity** which is characterized by an individual's physical differences, personal preferences, or life experiences.

**Cultural diversity** which is characterized by embracing differing beliefs, values and personal characteristics.

**Systems diversity** which is characterized by the organizational and management structures of our information institutions, systems and practices.

### 3. Identify EDIA issues in your AI project. (Think-pair-share)

What EDIA issues need to be considered and addressed in your AI project?

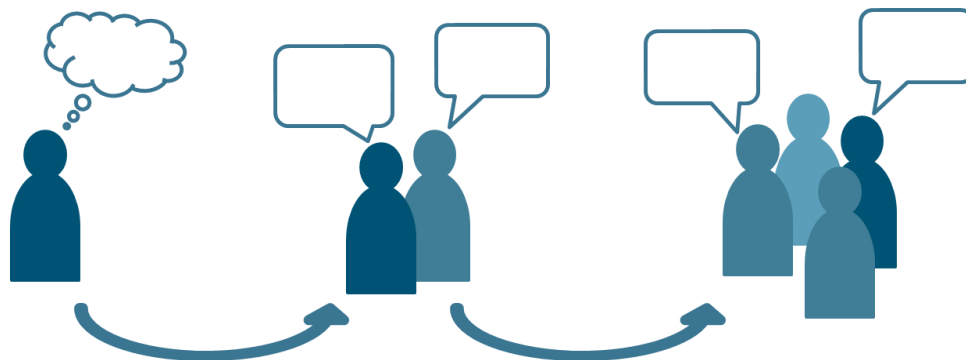


Image source: [https://canvas.csun.edu/courses/93131/pages/exponential-think-pair-share?module\\_item\\_id=3081401](https://canvas.csun.edu/courses/93131/pages/exponential-think-pair-share?module_item_id=3081401)

Issues: positionality, intersectionality; participants on decision-making; development and implementation; multiple ways of knowing; ethics; data availability and access; user considerations

### 4. Write your AI idea as a goal, using the [SMARTIE approach](#) (see Appendix A):

S.M.A.R.T.: Specific, Measurable, Achievable, Relevant, Time-Bound

to SMARTIE: Strategic, Measurable, Ambitious, Realistic, Time-Bound, Inclusive, Equitable

## References (selected):

### Definitions

- American Alliance of Museums. *Definitions of Diversity, Equity, Accessibility, and Inclusion*. <infographic>
- College of the Environment, University of Washington. *Diversity, Equity and Inclusion Glossary*. <https://environment.uw.edu/about/diversity-equity-inclusion/tools-and-additional-resources/glossary-dei-concepts/>
- Office of Diversity, Texas A&M University. *Glossary of Terms*. <https://diversity.tamu.edu/Menu/Glossary>

Intersectionality (see: <https://thirdsectorcompany.com/whom-does-deia-include/>) - finding common ground

### Diversity plan/goals

- Setting Diversity Goals: A Key to EDI Success <https://www.linkedin.com/pulse/setting-diversity-goals-key-edi-success-deetta-jones>

### Policies

- Pennsylvania Historical and Museum Commission. *Diversity, Equity, Inclusion, & Access Policy*. Adopted December 5, 2018 <https://www.phmc.pa.gov/About/Pages/Diversity-Policy-.aspx>

### Research

- EDI in Research: AI Biases as a Case Study, Souhila Baba | 13 Mar 2020 <https://www.mcgill.ca/equity/article/edi-research-ai-biases-case-study>

### Systemic Racism in Data Practices

Topics proposed for *IASSIST Quarterly issue on Systemic Racism in Data Practices*, co-edited by Jonathan Cain, Columbia University, and Trevor Watkins, George Mason University

- Implicit or explicit bias in AI and Machine Learning
- Data activism
- Anti-Indigeneity in big data
- Decolonizing data and data science
- Decolonizing scholarly data
- Bias in data collection practices
- Data and racial disparities in health sciences
- Race and precision medicine
- Racist practices in data reporting – Climate change, Covid-19, etc. reporting in marginalized communities
- Ethics of algorithm design
- Equity in data education – More inclusive (marginalized and underrepresented communities) -How do we improve this?
- Theories or suggestions on fair, ethical or trustworthy AI



## Conversational AI Part 1 - Theoretical

### Purpose:

This session will discuss the basic concepts, building blocks, interface design, techniques, and implementation of conversational AI. Specific topics are:

- Conversational user interfaces
- Main platforms for implementing chatbots
- Building blocks of chatbots
- Understanding natural language
- Integrating with voice interaction
- Digital assistant development (with DialogFlow)
- Conversational AI applications in libraries
- Social implications of conversational AI

**Date and/or Time:** Wednesday July 13, 8:30 AM – 10:30 AM

**Materials:** Slides, a DialogFlow example project, and an exercise instruction.

**Who:** Jiangen He, [jiangen@utk.edu](mailto:jiangen@utk.edu)

**How:** This session will use lectures, discussion, and a hands-on exercise.

### References:

DialogFlow Tutorials & samples <https://cloud.google.com/dialogflow/es/docs/tutorials>



**Linked Open Data (LOD)**  
**William Mischo, Timothy Cole, M.J. Han**

**Purpose (or learning goals):**

Discussion and demonstration of Linked Open Data technologies and activities over the last several years.

**Date and Time:** Wednesday July 13, 10:45 AM – 12:00 PM

**Materials:** PowerPoint Slides and Web browser demonstrations

**Who:** William H. Mischo, University of Illinois at Urbana-Champaign

**How:** Presentation, Demonstration, and Discussion

- Enhancing Identification, Selection, and Serendipity through *Linked Open Data*
- Discovery is not always a single step process
- Library users seek context and descriptive cues to help them understand, identify and select from the resources returned by a search
- Library systems want to incorporate systems like Google's Knowledge Cards and People Also Ask lists and answers
- Discovery of additional resources and follow-on searches through automated suggestions, through juxtaposition and inter-object relationships, and through serendipity
- Libraries as producers
- Libraries as consumers
- LD4L, LD4P, LD4P2, LD4P3
- LoC and BIBFRAME; OCLC and schema.org



## Machine Learning for Text with Topic Modeling and Clustering

### Purpose (or learning goals):

Discussion and demonstration of unsupervised clustering ML techniques and topic modeling

**Date and/or Time:** Wednesday July 13, 1:00 PM - 3:00 PM

**Materials:** PowerPoint Slides and Web browser demonstrations

**Who:** William H. Mischo, University of Illinois at Urbana-Champaign

**How:** Presentation, Demonstration, and Discussion

- Unsupervised ML clustering techniques
- Synergy between Bibliometric services and Machine Learning services in libraries
- Library data; processing, manipulating and cleaning
- Vectorizing techniques
- Vector similarity or distancing formulas
- Text as Bag of Words or Sequences (phrases)
- K-Means, K-Means++, Spectral Clustering, Custom
- Environments: ScikitLearn, MS Azure, Google Cloud, Amazon, Wolfram
- Biofuels example
- Cancer Center of Illinois example
- Clustering issues
- Phrase extraction algorithms

### References:

Qiankun Zhao, Ximing Cai, William H. Mischo, Liyuan Ma, “How do the Research and Public Communities view Biofuel Development”, *Renewable and Sustainable Energy Reviews*, November 2020, volume 133, article number 110265 <https://doi.org/10.1016/j.rser.2020.110265>

Aggarwal, Charu C., *Machine Learning for Text*, Springer, 2018.  
<https://doi.org/10.1007/978-3-319-75351-3>

[Machine Learning + Libraries Summit Event Summary](#)

[Machine Learning, Libraries, and Cross-Disciplinary Research: Possibilities and Provocations](#)

[Responsible Operations: Data Science, Machine Learning, and AI in Libraries](#)

[Machine Learning +Libraries: A Report on the State of the Field](#)



## Networking Dinner with Austin AI Experts

### Purpose (or learning goals):

- Connect Institute Fellows with AI experts in Austin who use AI in various capacities in research.
- Exchange AI project ideas, applications, issues, and challenges with AI experts in small groups.

**Date and Time:** Wednesday July 13, 5:30 PM - 7:30 PM

**Location:** Hilton Garden Inn, 17th Floor Apex Rooftop

**How:** Lightning talks by AI experts and exchange of ideas

### Who:

- Aaron Choate, Director of Research and Strategy, University of Texas Libraries.  
[//www.lib.utexas.edu/about/directory/aaron-choate](http://www.lib.utexas.edu/about/directory/aaron-choate)
- Ken Fleischmann, Professor, University of Texas at Austin School of Information and the Founding Chair of the Executive Team for Good Systems, a UT Grand Challenge.  
<https://www.ischool.utexas.edu/people/people-details?PersonID=220>
- Elliot Hauser, Assistant Professor, University of Texas at Austin School of Information.  
<https://elliotthouser.com/>
- Min Kyung Lee, Assistant Professor, University of Texas at Austin School of Information.  
<http://minlee.net/>
- Stephen Straus, Co-Founder and Managing Director, KUNG.AU.  
<https://www.kungfu.ai/person/stephen-straus>



## Conversational AI Part 2 - Application

### Purpose:

This session will introduce the key features and concepts of Google DialogFlow and show how it can be used to build chatbots for a variety of applications. Specific objectives of this session include:

- Creating intents
- Creating entities and parameters
- Adding follow-up intents
- Input and output context
- Creating a fulfillment
- Integrating a chatbot with your website

**Date and/or Time:** Thursday July 14, 8:30 AM - 11:45 AM

**Materials:** Slides, a DialogFlow example project, and an exercise instruction.

**Who:** Jiangen He, [jiangen@utk.edu](mailto:jiangen@utk.edu)

**How:** This session will use lectures, discussion, and a hands-on exercise.

### References:

DialogFlow Tutorials & samples <https://cloud.google.com/dialogflow/es/docs/tutorials>



## Post-Program Feedback and Activities

### Purpose (or learning goals):

- Learn of opportunities to demonstrate knowledge and application of AI library solutions.
- Consider modes to disseminate AI library applications and knowledge.
- Provide feedback at end of Institute and 6 months post-Institute

**Date and Time:** Thursday July 14, 11:45 AM – 12:00 PM

**Who:** Clara M. Chu, Director and Mortenson Distinguished Professor, Mortenson Center for International Library Programs, University of Illinois at Urbana-Champaign.

Email: [cmchu@illinois.edu](mailto:cmchu@illinois.edu)

### How:

#### 1. Learn about Post-Institute Opportunities with IDEA Institute on AI Partners: Association for Information Science and Technology (ASIS&T) and American Library Association (ALA).

1a. Present or attend at workshop of the 85th Annual Meeting of the Association for Information Science and Technology; October 29 - November 1, 2022 <https://www.asist.org/am22/>

SATURDAY, 29 OCTOBER 2022; 1:00-5:00 PM (see <https://www.asist.org/am22/am22-workshops/>)

*AI in the Real World: Strengthening Connections Between LIS Research and Practice (SIG-AI)*  
Soo Young Rieh, The University of Texas at Austin, USA; Clara M. Chu, University of Illinois at Urbana-Champaign, USA; Dania Bilal, University of Tennessee-Knoxville, USA

The workshop builds on and extends the 2021 workshop to strengthen the ASIS&T AI community by bringing together researchers, educators, students, and practitioners interested in designing AI applications and conducting AI research in the LIS field. This workshop covers a wide range of topics including AI applications, solutions, empirical research findings, and perspectives in library and information environments. The participants will discuss lessons learned from designing AI applications and solutions in the context of library and information environments and work together to brainstorm potential solutions for making AI more effective and transferable for library users. Through the World Café method (i.e., collaborative dialogue with rotating multiple breakout sessions that build on each other so that issues are considered in-depth) and plenary discussion, participants will generate advanced AI solutions and research agendas for future investigations.

1b. Participate in [Core: Leadership, Infrastructure, Futures](#) (a division of the American Library Association) events.

Mission: To cultivate and amplify the collective expertise of library workers in core functions through community building, advocacy, and learning. Join the ALA CORE Artificial Intelligence and Machine Learning in Libraries Interest Group - no need to join ALA.

<https://www.ala.org/core/member-center/interest-groups/artificial-intelligence-and-machine-learning-in-libraries>





## 2. Explore dissemination modes. (group discussion)

- conference participation
- article publication
- workplace presentation
- other

## 3. Provide feedback at end of Institute and 6 months post-Institute

- Fellows: Post-Institute Survey (due: Monday, July 25th) and 6-month post-Institute Survey
- Fellows' supervisors: 6-month post-Institute Survey



**2022 Fellows AI Project and Action Plan**

Name: \_\_\_\_\_ Email: \_\_\_\_\_

1. What problem/need are you addressing?
  
2. Which user group will benefit from this idea?
  
3. What are the desired results (measurable change) and outcome (future state of being)?
  
4. Title of your project and description (max. 150 words)
  
5. Do you need to get permission from a supervisor, director, and/or colleagues before implementing the goal? If yes, state who it is and explain their role.

6. **Write your AI idea as a goal, using the SMARTIE approach:**  
 S.M.A.R.T.: Specific, Measurable, Achievable, Relevant, Time-Bound  
 to SMARTIE: Strategic, Measurable, Ambitious, Realistic, Time-Bound, Inclusive, Equitable

**7. Your Action Plan:**

<i>Actions</i>	<i>Timelines</i>	<i>Responsibility of</i>	<i>Partners</i>
1.			
2.			
3.			
4.			
5.			
6.			
7.			
8. <add more rows as needed>			

8. Do you need any resources or funding to implement this Action Plan? If so, list them.

9. How will you measure the impact of this action?

10. Comments



## Capstone Showcase Guidelines

### Purpose (or learning goals)

- Sharing learning on AI gained from the Institute on AI related to Fellows' AI projects
- Introduce AI projects in library and information environments to local library and information professionals and stakeholders.

### Date and Time

Friday July 15, 9:00 AM – 12:00 PM

### Materials:

Complete and upload to <https://uofi.app.box.com/f/9172abf98f6d413bb641f472a7dd12b0> by Thursday, July 14, 2022; 7 pm CT the following documents (available at [go.illinois.edu/ideainstitute2022](http://go.illinois.edu/ideainstitute2022)):

- Project and Action Plan template
- Presentation template (PPT)

Moderator: Clara M. Chu,

Presenters: IDEA Institute on AI Fellows

Audience: The capstone showcase is open to local library and information professionals and stakeholders. Due to COVID-19 precautions, only selected invitations will be made. The session will be hybrid: in person and on Zoom. Please invite your supervisors and colleges if you like using this Zoom link:

<https://utexas.zoom.us/j/98704306212?pwd=R2pVd1VMVWVQekowM0pOSTNhcFdRZz09>

Meeting ID: 987 0430 6212 Passcode: 404857

### How:

Each Fellow will do a 5-minute presentation (3 content slides max.) in a 4 or 5-person panel, followed by a 8-minute Q&A panel session. The presentation order will be provided by Thursday, July 14th.

9:00-9:10 Welcome

9:10-9:40 Panel 1

9:40-10:10 Panel 2

10:10-10:20 Stretch Break

10:20-10:50 Panel 3

10:50-11:30 Panel 4

11:30-11:45 Certificate Presentation

11:45-12:00 Final Remarks, Photo and Farewell

## Capstone: Appendix A

# SMARTIE Goals Worksheet

Source: [The Management Center](#) <[Google worksheet](#)>

Goals are a concrete way to drive results, but without an explicit equity and inclusion component, goals won't produce better outcomes for marginalized communities, address disparities, or support belonging. Introducing SMARTIE goals! SMARTIE stands for:

### STRATEGIC

Reflects an important dimension of what your organization seeks to accomplish (programmatic or capacity-building priorities).

### MEASURABLE

Includes standards by which reasonable people can agree on whether the goal has been met (by numbers or defined qualities).

### AMBITIOUS

Challenging enough that achievement would mean significant progress—a “stretch” for the organization.

### REALISTIC

Not so challenging as to indicate lack of thought about resources, capacity, or execution; possible to track and worth the time and energy to do so.

### TIME-BOUND

Includes a clear deadline.

### INCLUSIVE

Brings traditionally marginalized people—particularly those most impacted—into processes, activities, and decision/policy-making in a way that shares power. [NOTE: Differentiate between inclusion vs. tokenism. For AI project, also consider inclusivity in the actual content/project/solution.]

### EQUITABLE

Seeks to address systemic injustice, inequity, or oppression.

By incorporating equity and inclusion into your SMART goals, you can make sure your organization's commitment to racial equity and inclusion is anchored by tangible and actionable steps. Here's an example of a SMART goal turned SMARTIE:

SMART	SMARTIE
Build a volunteer team of 100 door-to-door canvassers by May...	...with at least 10 people of color recruited as volunteer leaders first, so that they can help shape the way we run the canvasses.

#### ====Resources====

See other examples: <https://www.idealists.org/en/careers/better-than-smart-smartie-goals>

On action plans: Sonnie, A. (2018). [Advancing Racial Equity in Public Libraries: Case Studies from the Field](#) pp. 35-36). Government Alliance on Race and Equity (GARE).

## Start Writing Your SMARTIE Goals

Use this template to write a goal for yourself or a team member.

**Time Bound:** My goals between [ ] (start date) and [ ] (end date) are to achieve this **Strategic** and **Ambitious** outcome:

I will know success when I see it using these **Measurable** standards:

- 
- 

A **Realistic** plan to achieve this goal includes these tactics/activities (consider time, resources, capacity):

By [ ] (date)

By [ ] (date)

By [ ] (date)

Thinking about **Equity and Inclusion**: Can you imagine there being any unintentional *disparate impact* along lines of power and identity? How might inequity or exclusion show up? For whom?

How could you *change the goal* to either mitigate that disparate impact or make **Equity and Inclusion** more explicit?

## Capstone: Appendix B

### An action plan includes the following components:

Source: Sonnie, A. (2018). [Advancing Racial Equity in Public Libraries: Case Studies from the Field](#) pp. 35-36). Government Alliance on Race and Equity (GARE).

**Desired Results:** Community-level conditions you aim to achieve (the change you want to see).

**Community Indicator:** The means by which you can measure improved conditions.

**Outcome:** A future state of being, resulting from a change at the jurisdiction, department or program level. Strong outcomes articulate a clear improvement or define how much improvement will take place.

**Actions:** Specific things your library will do to achieve the outcome.

**Performance Measure:** A quantifiable measure of how well an action is working.

Different types of measures include:

- A. **Quantity:** How much did we do?
- B. **Quality:** How well did we do it?
- C. **Impact:** Is anyone better off?

**Timeline:** The month, quarter, and/or year(s) an action will be accomplished.

**Accountability:** The position or body responsible for the action and/or accountable for its completion.

[Plus funding/resources]



## Information Sheet

July 10-15, 2022; Austin, TX @ School of Information

This sheet provides information to help you to prepare for and participate in the IDEA Institute on AI. Additional information can be found at <http://idea.infosci.utk.edu>

In case of **emergency**, text or call **Soo Young Rieh** at (734) 330-6431 or email at [rieh@ischool.utexas.edu](mailto:rieh@ischool.utexas.edu).

### Fellows Get-Together

All Fellows are encouraged to reach out to each other to connect on Saturday. Use the [ideainstitute22@utlists.utexas.edu](mailto:ideainstitute22@utlists.utexas.edu) list to do so.

### Face Masks ([UT Austin Guidance](#))

UT community members in higher risk environments are expected to use face coverings with increased protective ability. In general, cloth face masks are not appropriate substitutes for surgical masks or respirators in workplaces where the latter are recommended and available.

### Maps

The School of Information, also known as the UT Administration Building, is located at [1616 Guadalupe St. Suite #5.202, Austin, TX 78701](#).

The University of Texas at Austin campus map can be found [Main Campus Map](#).

Directions from the Institute hotel (Hilton Garden Inn - Austin University Capitol District) to the School of Information can be found [Google Map](#).

### Institute Program

The institute's program features six days, consisting of welcome and introduction, technical sessions, lunch discussion sessions, networking discussion dinners, AI project idea exchange, AI project activities, and a capstone showcase.

### Lodging

Reservations have been made for Fellows and Advisory Board members who are staying at the **Hilton Garden Inn - Austin University Capitol District**, located at 301 W 17th St, Austin, TX 78701; phone: 512-319-3333.



## Internet Connection

Access to [Wi-Fi internet connection for visitors](#) during the Institute is provided through the University of Texas at Austin. Fellows may use *utguest* or *eduroam* to connect to the network. To use *utguest*, simply select the network on your device and follow the instructions here: <https://www.utguest.org/>

## Fellowship

The funded Fellowship includes up to \$500 in airfare, or in mileage if driving is preferred. Mileage is reimbursed at \$0.585/per mile. Provide a receipt for the airfare as applicable. If you drive, capture a screenshot of the map for your trip showing the departure and return cities and distance. In addition, the Fellowship covers up to \$100 in local transportation.

Parking at the hotel is \$33.75/day and the nearby School of Information Parking is \$18/day (1616 Guadalupe St., Austin, TX). You do not need a permit to park in the School of Information Parking Structure. Funded Fellows may claim the parking cost as part of the \$100 in transportation. A receipt is required for reimbursement.

## Meals and Per Diem

Meals covered during the Institute include **6 lunches** and **2 dinners**. *No breakfast* will be provided during the Institute. **Hilton Garden Inn** provides a light complimentary breakfast for free.

Per diem. Meals outside the Institute are reimbursed as per diem according to the UT-Austin guidelines: <https://wikis.utexas.edu/display/embusinessservices/Meals+and+Lodging+Per+Diem+Limits>. No receipts are required for these meals. Since Hilton Garden Inn provides a complimentary light breakfast, the per diem will exclude breakfast.

## Reimbursement

Funded Fellows, Advisory Board members, and Team, submit your receipts to Jennifer Kryska ([jennifer.kryska@ischool.utexas.edu](mailto:jennifer.kryska@ischool.utexas.edu)) for reimbursement **no later than July 29th**.

Keep copies of the original receipts for airfare, local transportation, and parking. It may take 4-6 weeks to receive your reimbursement check. The check will be sent to the address you provided in the travel form.

## Institute Location

The Institute is held in **The School of Information**, also known as the UT Administration Building, located at **1616 Guadalupe St, Suite #5.222, Austin, TX 78701**.

Arrival: The program will start at 9:00 AM on Sunday, July 10<sup>th</sup>. Arrive 15 minutes in advance for checking-in.

Meetings/Sessions: The Institute's meetings/sessions will be in the School of Information 5th floor, **Room 5.522**.

## Dress Code

The dress code is casual and/or semi-formal. Have a jacket handy, in case it gets cold in the building.

## Accessibility

A list of assistive technology is at <https://idea.infosci.utk.edu/>.

## Networking Dinners

Two networking dinners are scheduled during the Institute, one on Monday, July 11 at 5:30-8:00 PM and another on Wednesday, July 13 at 5:30-7:30 PM. On Monday, Jason Griffey is the guest speaker during dinner. On Wednesday, AI experts in Austin are guest speakers during dinner. Additional information about the guest speakers and their presentations will be provided in the detailed program.

## Film During the Institute

We will show the film, **Coded Bias**, on Tuesday, July 12<sup>th</sup> at 4:30-6:00 PM. Viewing the film is optional.

## Showcase

A showcase is scheduled on Friday, July 15<sup>th</sup> from 9 AM -12 PM to present your AI projects. Local information professionals will be invited to attend the showcase. Materials to prepare for the showcase and for developing your AI projects are found in the [Box folder](#).

## Blogging

Your blogging entry should be submitted by the end of the Institute, Friday, July 15<sup>th</sup>. The blogging instructions can be found in the [Box folder](#).

## Certificate of Institute Completion

You will receive a certificate of Institute completion on Friday, July 15<sup>th</sup> after the showcase.

## Tweeting

You are encouraged to tweet about the Institute using these hashtags:

**#IDEA\_AI\_Institute**

**#Allibraries**

## Restaurants in Austin

A list of restaurants in Austin is at <https://www.austintexas.org/things-to-do/food-and-drink/>

## Attractions in Austin

A list of top attractions in Austin is at <https://www.austintexas.org/>

## Follow Up after the Institute

Six months after the Institute, we plan to follow-up with you about using AI in the workplace. We will

email you a link to a 6-month-post-Institute survey to reflect on your AI application in the workplace and schedule a meeting.