iSchool Research Showcase 2023

Poster Abstracts

Activist organizations and their strategies to influence the legalization of medical cannabis in Brazil

Janaynne Carvalho do Amaral, Jodi Schneider

Little is known about whether or how argumentative strategies have changed with the rise in online activism in health. Recent work on activist argumentation online examines fashion sustainability (Mercuri, 2023) and environmental risks (Brambilla, 2020). In 2015, after the mobilization of activists, Brazil legalized one form of medical cannabis by prescription only. In 2016, the Medical Cannabis Research and Patient Support Association (APEPI) registered as a Brazilian non-governmental organization focusing on “awareness, demystification and democratization of the medical use of cannabis” (APEPI, n.d.). Our project is a case study using netnography and discourse analysis to explore materials from APEPI’s early cannabis activism such as news, websites, interviews with the activists, and documentaries as well as the organization’s recent email newsletters. We analyze the visual and textual argumentative strategies that online cannabis activists have used and compare them to strategies used by AIDS activists in the 1980’s, before the rise in online communication: (1) redefining the legitimate participants in discourse about the disease by questioning the authority of doctors and researchers and (2) translating scientific information for patients while sharing patients’ needs with doctors and researchers (Fabj & Sobnosky, 2017). We hypothesize that online cannabis activists are using the same argumentative strategies adopted by AIDS activists.
Keywords: Online activism, Dissemination of knowledge, Visual Argumentation, Webdesign

Cause and Effects Between COVID-19 Cases and Related Social Media and Google Search Activity

Thierry Guigma

The COVID-19 pandemic has not only posed a severe global health threat but has also transformed the way information is shared and consumed. This study aims to uncover the intricate cause-and-effect relationships between the pandemic's progression and digital interactions. Using impulse response analysis, typically employed in signal processing and econometrics, the research explores how COVID-19 cases and online activity influence each other, shedding light on the underlying societal and behavioral factors.The findings reveal a significant one-way impact, where an increase in COVID-19 cases leads to an immediate rise in COVID-19-related tweets and Google searches about symptoms. This demonstrates that as the pandemic evolves, public interest and engagement with COVID-19 topics intensify, driving online discussions and information-seeking behavior. Conversely, the reverse relationship, where online discussions influence COVID-19 cases, is weak or insignificant. This emphasizes that social media conversations and online searches do not have a direct and important impact on virus transmission.These results underscore the complex nature of

information dissemination during a pandemic. Real-world events and changes in the pandemic landscape have a more immediate effect on public communication and information-seeking behaviors than the ability of online discussions to directly shape the pandemic's trajectory. In conclusion, COVID-19 cases significantly influence the volume of COVID-19-related online content, highlighting the dynamic interplay between the pandemic's progression and digital communication.
Keywords: COVID-19, Twitter, Google Search, Impulse Response Analysis,

Coconut Library Tool: A Web-Based Application for Advanced Textual Analysis for Librarians

Manika Lamba, Faizhal Arif Santosa and Crissandra J. George

In this era where information explosion is inevitable, librarians are compelled to adopt a novel methodology for executing textual analysis. This method requires sufficient knowledge of statistics and programming skills, which is still a challenge for librarians with a humanities or social sciences background. Several practices for using data science in libraries have been widely implemented, but these cannot necessarily be implemented easily in various libraries due to diverse library types, varying resources, and staff expertise. The Coconut Library Tool is a web-based application that employs multiple algorithmic NLP techniques at the backend, including topic modeling, lemmatization, stemming, and network analysis & visualization. For topic modeling, we have incorporated three algorithms, LDA, Biterm Topic Model, and BERTopic, allowing users to select their preferred method. By employing the Hugging Face space, the present application is capable of functioning without imposing any significant load on the user. Additionally, the application offers seamless analysis of custom CSV files and data collected from diverse indexing databases including Scopus, Web of Science, and Lens. Using this tool, librarians can perform (i) topic modeling for making ontologies, automatic subject classification, recommendation services, bibliometrics, altmetrics, and better resource searching and retrieval; (ii) network text analysis to improve knowledge discovery, obtain deeper insights, and support scholars meaningfully, ultimately enhancing the library's services and resources; (iii) sunburst visualization to easily browse and grasp the relationships between different levels of the hierarchy; and (iv) text pre-processing which includes lemmatization & stemming for locating basic words and aids in catching the true meaning of the word, which can lead to improved semantic analysis and comprehension of the text. The primary aim of this application is to directly support librarians in delivering high-quality services, promoting sustainable collaboration among diverse stakeholders, and demonstrating the inclusive and equitable use of artificial intelligence tools by librarians across the board.
Keywords: Natural Language Processing, Topic Modeling, Network Visualization, Network Analysis, Textual Analysis

Cogent Constructs: Uniting Cognitive Theory with Generative AI in Multi-Agent Systems

Tre Tomaszewski

This research extends the idea of "generative agents" by integrating cognitive systems modeled after frameworks such as ACT-R, Soar, and CLARION, with the intent to produce multi-agent systems with enhanced capabilities and the ability to communicate using natural language. In their original formulation, found in “Generative Agents” (Park et al, 2023), generative agents employ generative AI/ML, specifically LLMs, to create mimicries of human behavior and facilitating inter-agent interactions. A key feature is the information propagation between agents through conversation. Additionally, these agents can plan and “introspect” using LLM alongside textual perceptions and associative embeddings. However, the cognitive loop implementation is simplistic and highly reliant on OpenAI’s proprietary GPT. As details such as system architecture are not publicly accessible, appropriate, and scientific assessment of the agents’ and their abilities is impossible.This in-progress study aims to assess the efficacy of agents equipped with specific modules derived from cognitive theory such as declarative memory and procedural memory, goal formulation, and meta-cognitive introspection. These will enable continual learning from both experience and intra-agent or human instruction, while the known effectiveness of the frameworks will mediate the reliance on the LLM.To fully realize their ability, the agents will be embedded in an interactive virtual environment alongside other agents, each seeded with idiosyncratic knowledge and behaviors. The emergent capabilities of these agents will be compared qualitatively to the original generative agents and quantitatively against the agent models implemented using raw cognitive frameworks. Additionally, the research explores the potential of these systems in improving agents that use the less capable, open LLMs, like Meta’s "Llama".Potential broader impacts of successful outcomes include: analysis of simulated social phenomena and information dissemination, enable "online learning" alongside generative models, enhancement of human-agent collaboration, improved scrutiny of LLM information flow for alignment/safety, and advancements in explainable AI.
Keywords: agents, cognitive theory, generative ai, social simulation

Connection, Context, and Content: Adult Leaders’ Approaches to Out-of-School Time Youth Development and Support

Rachel Magee, Amy Leman, Kate Suchodolski

Youth involvement in out-of-school time (OST) programs is consistently linked to positive developmental outcomes in all youth and specifically underrepresented youth. Within OST programs, relationships developed with adult staff are important for positive experiences but not all OST programs promote the same perceptions of supportive relationships. For this poster, we aim to understand how adults in OST programs provide spaces that support positive youth outcomes. This is part of a larger study conducted in 2020-2022 that examined the practices and experiences of OST providers and youth participants across 5 community programs in a mid-sized, midwestern city. We analyzed interviews with 13 staff to understand

their approaches and process and identified three environmental elements that enable OST staff to support youth developmental needs: Content, Context, and Connection. Content is the materials, activities, and processes built into programs to support youth developmental outcomes. Context describes the lived experiences youth and adults in programs bring to their interactions as well as the adults' awareness of youth developmental stages, experiences, and community environments. Connection, is built upon recognition and honoring that youth are individuals deserving of respect and is co-created by invested adults and youth willing to be vulnerable and extend trust. Connection is needed in order to be able to accomplish the goals and objectives addressed by content and only happens when adults acknowledge the contexts in which youth arrive at the OST program.In OST programs, even though developmental relationships are known to be important, and OST staff express the importance of spending time on connection to engage youth, time and resources are focused first on content or activities for skill and knowledge growth. Content drives financial support. Our research is the start of a conversation to extend OST program priorities beyond content goals to value staff time for connection and relationship development.
Keywords: youth, teens, positive development, relationships, out-of-school time

Dams of Time: Navigating Ethics in Geospatial Data Science

Ivan Kong, Peter Darch

Researchers often struggle with ethical decisions when using geospatial data due to complex social and technical factors. Understanding their work practices is crucial for effective interventions such as tools, checklists, and training.Here, we present the results of a case study of researchers using an emerging geospatial data and AI system, the NSF-funded Institute for Geospatial Understanding through an Integrated Discovery Environment (I-GUIDE), to study the risks of aging dam infrastructure in the US. These researchers have multidisciplinary science and social science backgrounds and exhibit high methodological diversity. They integrate data from a range of publicly-accessible sources to predict the possible failure of aging dams and to offer solutions to stakeholders to manage the effects of dam failure. We have conducted interviews (n=16) and observation work with researchers over a period of 6 months; the study is still ongoing.Relevant ethical issues facing aging dam researchers include privacy and consent of data subjects, bias in datasets, data sovereignty, and supporting users of their work products. Many employ a diverse range of strategies to address ethical issues, including inspecting data for bias; practicing informed consent with interview subjects; making data transparent and available; using data only from reputable sources; and taking steps to improve data quality. However, researchers struggle in the face of pressure to produce publishable results. Problems multiply when researchers use data they did not collect themselves, when datasets are transferred across institutional, disciplinary, and international boundaries as they may be subject to divergent policies and understandings of what is ethical.Our study demonstrates researchers have a general understanding of producing ethically sound research but can benefit from processes that will further strengthen their research decision-making while placing little additional burden

on already fully-stretched researchers. Examples of processes include ethical decision-making frameworks; comprehensive data ethics checklists and guides; transparent methodologies; and engaging domain experts in the research endeavor.
Keywords: geoscience data science; ethics; IGUIDE

Disinformation Tactics and Designing to Deceive

Emily Wegrzyn

Disinformation is an age-old problem with modern day upgrades. Today's disinformation operations seek to influence the attitudes, minds, and behaviors of targeted populations and individuals. 20th Century conflicts systematized it into military and state crafts. By the end of World War II, most world powers employed it to various extents with global reach. Advances in electronic communications combined with increased globalization accelerate disinformation's evolution. Its application expands beyond the military and political into ideological areas including society, education, medicine, and culture. Agents of influence are influential people with power, position, or credibility who can promote specific objectives. Modern examples of prominent agents of influence are Venezuelan President, Nicolás Maduro (Shuya 2019, 25), and perennial French Presidential candidate Marine LePen (LeBourg 2018, 2, 17). Mixing in false information alongside carefully curated true information is a technique used to simulate legitimacy. The 2020 American Presidential Election was fraught with inaccurate assertions of voter fraud. Real videos of a ballot delivery in Detroit (Swenson 2021) and ballot storage cases in Atlanta (Brown 2020) buttressed false accusations in news and social media.Saturation introduces disinformation in fringe outlets, then propagates it through additional channels working it into mainstream consciousness. An interview of Dr. Michael Yeadon, former vice-president of Pfizer Inc., the company that collaborated with BioNtech on the first COVID-19 vaccination, aired on the German OVALmedia website in late 2020 (OVALmedia 2020). Soon after, the anti-mask, anti-vaccine, anti-distancing diatribe made its way across Europe and around the world. Dr. Yeadon's statements have been disproven many times over (Stecklow and MacCaskill 2021), but the video persists.Disinformation operations have many targets. Some of these are ideological enemies, policies, or social, scientific, and technical achievements. Tactics used undermine confidence, discredit trusted institutions, and advance conditions favorable for achieving objectives. They are designed to deceive.
Keywords: Disinformation, information operations, agents of influence, infodemic, falsehoods

Enhancing SemRep Coverage Using a Relation Classification Approach

Shufan Ming

Relation extraction is an essential task in the field of biomedical natural language processing and offers significant benefits to various downstream applications, including drug repurposing, literature-based discovery, and knowledge graph construction. The rule-based system SemRep, developed at the U.S. National Library of Medicine, has established a solid

baseline for extracting subject-predicate-object triplets from unstructured biomedical text. SemRep serves as the backbone for constructing the Semantic MEDLINE Database (SemMedDB), which stores semantic relations between biomedical terms mined from the entire PubMed literature. In this study, we leveraged the capabilities of the pre-trained PubMedBERT model, enhancing it through additional contrastive pre-training and fine-tuning on our human-curated task-specific dataset comprising 23 relation types. Our goal was to discover more relations to improve SemRep's recall and expand the size of SemMedDB. We evaluated the performance of our PubMedBERT-based model on the SemRep Gold Standard dataset, revealing that the model yields an average of 37% more recall than SemRep. Additionally, we conducted a large-scale assessment, testing the model on 12,000 abstracts randomly selected from PubMed. We manually assessed the quality of 500 relations predicted by the model that SemRep had not previously identified, and the model achieved an estimated precision of 65% on these relations. These findings underscore the promise of our model in achieving more comprehensive coverage of relationships mentioned in biomedical texts, thereby enhancing various downstream applications in the field.

Keywords: Biomedical relation extraction, Large language model, SemRep

Epistemological responsibility in law and science: Sharing the burden

Ted Ledford

Professional standards differ between scientists and litigators when judging the explanatory use of evidence. Yet, scientists have a role as expert witnesses in legal cases to assess the strength of evidence and present new facts. How evidence can support a claim differs in legal argument versus a scientific interpretation. The Federal Rules for Evidence handle the admissible guidelines for what should count as scientific evidence and eligibility for scientific witness expertise in court. According to the Frye standard (1923), acceptability depends on established credibility in a scientific community; it was updated later (ratified in some states) by the Daubert series (1994), which emphasizes reliability of methodology and relevance of expertise bearing to the legal facts of a case, using a judge’s discretion. In the transition from Frye to Daubert, how must litigators adjust arguments and seek alternative evidence from scientific sources? Under Frye, a scientist’s credibility to scrutinize evidence becomes entrenched in the genealogy of specialties, methodologies and practices dividing how scientists are qualified to make inferences across domains. And yet, under Daubert, a litigation team can conduct scientific inquiry under the guise of fact-building by recruiting disparate expert scientists, with a range of backgrounds, to draw from a combination of bodies of evidence that bears relevance only by methodological analogy or similarity pertaining to the case. When Daubert was introduced, did scientists persist as expert witnesses in similar cases and, when funded by litigators, become more likely to pursue research designed to produce evidence favorable toward the new admissible guidelines in court? This research will study the precedent for how litigators, on behalf of industry (Big

Tobacco), selected evidence to construct scientific claims, and how the utility of scientific information under the law can improve.
Keywords: evidence, litigation, epistemology, philosophy of science

Graph Neural Bandits

Yunzhe Qi

Contextual bandits aim to choose the optimal arm with the highest reward out of a set of candidates based on their contextual information, and various bandit algorithms have been applied to real-world applications due to their ability of solving the exploitation-exploration dilemma. Motivated by online recommendation scenarios, in this paper, we propose a framework named Graph Neural Bandits (GNB) to leverage the collaborative nature among users empowered by graph neural networks (GNNs). Instead of estimating rigid user clusters as in existing works, we model the "fine-grained" collaborative effects through estimated user graphs in terms of exploitation and exploration respectively. Then, to refine the recommendation strategy, we utilize separate GNN-based models on estimated user graphs for exploitation and adaptive exploration. Theoretical analysis and experimental results on multiple real data sets in comparison with state-of-the-art baselines are provided to demonstrate the effectiveness of our proposed framework.
Keywords: Contextual Bandit; User Personalization

Harold Baron Digital Archival Research and Publication Project

Divya Pathak

This poster presents early-stage research on the Harold Baron Digital Archival Research and Publication Project. The project seeks to build an online library to showcase the life, work, and records of Harold (Hal) Baron, “(1930-2017), former Director of Research for the Chicago Urban League (1962-1968) and, later, public policy architect of the historic mayoral campaign of Harold Washington, followed by a tenure as the Chief Policy Advisor for the Washington Administration” (from https://halbaronproject.web.illinois.edu/omeka/). This poster will illustrate digital analysis conducted on the text of his unpublished manuscript, “An American Dilemma: Reconstructing the Rhetoric of Race,” and its accompanying notes and citations. These analyses include reconciling citations and references between the manuscript and the physical sources and notes, text cleaning of the digitized manuscript chapters, and visualizations of the content of the chapters using the digital text analysis environment, Voyant Tools.“Baron’s contributions have had an impact on the theory, practice and social policies of American race relations for over 50 years, contributions that he either originated or was an early proponent. His final, unfinished, work, An American Dilemma: Reconstructing the Rhetoric of Race will instigate a seismic shift in liberal social and policy science, making his voluminous papers from this project the capstone of his years of quiet, intense, and consequential effort to deconstruct the dominance of America’s racial formations” (from https://halbaronproject.web.illinois.edu/omeka/).

Keywords: Digital humanities, digital archives, civil rights, social policy, race relations

How Knowledge Intermediaries Gather and Make Sense of COVID-19 Information: An Interview Study

Togzhan Seilkhanova, Dr. Jodi Schneider

Our three-year project is planning case studies on three topics: COVID-19, AI and labor, and climate change. For each case study, we are both interviewing intermediaries and analyzing public-facing digital communications. We will apply findings from our research to public library service development, to co-develop services for intermediaries as well as improve librarians' comfort with helping the general public make sense of scientific and technical information.To date, our IRB-approved study has conducted and analyzed 12 semi-structured interviews with COVID-19 information intermediaries: advocacy group representatives, journalists, fact-checkers, filmmakers, lawyers, and university professors. Our interviews use the critical incident technique (Flanagan, 1954). First, we prompt participants to recall a specific incident in which they were gathering information. Then we ask them to describe their process.After transcribing interviews, we analyze them in MAXQDA using thematic analysis (Braun & Clarke, 2006) and modified grounded theory (Alnsour, 2022).Initial results show that our knowledge intermediaries follow the same general 4-step information process:1.Gather new information2.Determine the relevance3.Verify information4.Share with the audienceWe also find a pattern in how knowledge intermediaries and their audiences are formed: either an audience starts forming around someone who shares information, or knowledge intermediaries recognize an existing audience and offer information services targeted specifically at this audience. In addition, knowledge intermediaries can be grouped by their goals and motivations for sharing information: they can be more information-centered or more people-centered.

Keywords: knowledge intermediaries, information behavior, information sharing,

Incorporating Indigeneity: my goal for librarianship

Katherine Witzig

This work will focus on my efforts thus far and those in the foreseeable future to incorporate my Indigenous identity into professional librarianship. I will use existing scholarship and community knowledge, blended with my personal experiences, to propose a culturally responsive shift to existing professional values.

Keywords: Inclusivity; Indigenous identity; librarian values; professional values

A Large Language Model Based Approach to Detecting the Influence of Darwinian Ideas

Lucian Li

Scholars across several fields have attempted to study dynamics of intellectual influence. Cordell and Smith's Viral Texts project focuses on direct copying and text reuse. Barron (2018) et. al's study of the French Revolution relies on comparing topic model distributions between speeches to quantify influence. Finally, Soni and Klein's (2021) study of abolitionist newspapers uses change in word embeddings to speculate on the origin of new ideas. The unique affordances of LLMs in text vectorization enable more granular consideration of the mechanisms of knowledge transmission. LLMs have demonstrated strong performance on quantifying semantic similarity and detecting paraphrasing. (Becker 2023, https://arxiv.org/abs/2303.13989) LLMs may also be able to detect more abstract similarities in argumentation (Opitz 2021 https://doi.org/10.18653/v1/2021.argmining-1.3), potentially helpful in cases of more indirect influence rather than citation. An LLM based approach can provide confidence intervals, ranging from direct quotation, to paraphrasing, and finally speculative discovery of indirect influence. I constructed a nonfiction corpus based on authors referenced in influential academic periodicals. These citations capture a substantial subset of authors active in 19th century intellectual networks. I have constructed a corpus of roughly 200,000 works by these authors available on HathiTrust and the Internet Archive. Using these sources, I will conduct a proof of concept study to demonstrate the potential of LLMs for detecting influence. Focusing on Darwin's Origin of Species and Descent of Man, I will quantify their influence on scholarly discourses. Did they borrow more from previously published books (i. e. Crystallize ideas already in the public consciousness) or exert more influence on future works (i. e. Introduce generally new ideas into the scholarly discourse)? How does this "novelty" or "transience" compare to other works in the corpus?
Keywords: NLP, Digital Humanities, Computational Social Science, Bibliometrics

PREPP-ing for Practice: An Experiential Learning Experiment in Practicum and Capstone Design

Sharon L. Comstock, Maria Bonn, Ben Schaap

iSchool has a long history of experiential learning praxis grounded in inquiry- and problem-based learning, community informatics, and engaged scholarship pedagogies. The PRoject to Examine Professional Preparedness (PREPP), funded by University of Illinois, Urbana-Champaign Provost's Initiative on Teaching Advancement ($7,500) asks, "In what ways do our Practicum Experiences meet Program Learning Outcomes and needs of our Students entering the workforce?" This Project seeks to glean measurement data to place existing data proxies in meaningful contexts. (Bonn, Comstock, & Schaap, 2023 "Standardizing MSLIS Experiential Learning Opportunities": https://ischool.illinois.edu/news-events/news/2023/08/ischool-team-receives-funding-provost-study-experiential-learning). PREPP seeks to capture learner-centered strategies that bridge the artificiality of the classroom to the real-world of practice in meaningful and measurable ways. At present, the MSLIS program does not have a formal degree requirement for a culminating experience that has students apply their knowledge within professional contexts. While practicums and internships are encouraged, only some students opt into taking them. PREPP seeks to move

the program closer to identifying two aspects: (1) What kinds of culminating experiences benefit students and in what ways; and (2) How do these specific culminating experiences align with our MSLIS Program Learning Outcomes? To answer these, PREPP is using a modified experimental design, developing a pilot course where a sample of students would choose from course options in the final year of their program: (a) a traditional practicum and/or (b) a project-based learning (capstone) course. Since a notable number of MSLIS students are enrolled online, PREPP is specifically designed to cultivate equitable experiential learning opportunities for both on-campus and online modalities.
Keywords: Practicum, Experiential Learning, Workforce Development, Library & Information Sciences, MSLIS-CAS, Assessment

Quantifying the Impact of Scientific Documentaries Using Natural Language Processing

Sri Nithya Yeragorls

My research focuses on studying the impact of scientific documentaries on its viewers. One way in which I do this is with impact-analysis which is a natural language processing (NLP) technique applied to reviews left on the documentary pages of platforms like Amazon and YouTube. In this research that technique consists of extracting reviews from the platforms mentioned above and annotating them using impact categories we have created. Later these annotated sentences will be used for training NLP models. One of the main findings from the research is that most viewers make positive comments about documentaries after watching them.
Keywords: Scientific Documentaries, NLP, Impact-Analysis

Research with User-Generated Book Review Data: Legal and Ethical Pitfalls and Contextualized Mitigations

Yuerong Hu, Glen Layne-Worthey, Alaine Martaus, J. Stephen Downie, Jana Diesner

The growing quantity of user-generated book reviews has opened up unprecedented opportunities for empirical research on books, reading, and readership. While there is an abundance of literature addressing the legal and ethical use of user-generated and social media data in general, for user-generated book reviews, such discussions have been mostly absent. From a library and information sciences perspective, user-generated book reviews can pose novel challenges because each book reviewer may simultaneously be (1) a presumably anonymous and safe online user; and, (2) an identifiable reader who can suffer real harm, e.g., cyber doxing and personal attack. This user/reader duality can create conflicting recommendations regarding which legal or ethical guidelines to follow. According to our review, potential legal issues include copyright infringement and violations of terms of service/end-user license agreements and privacy rights, while ethical concerns are centered on users' expectations, informed consent, and institutional reviews. This paper

reviews (1) potential legal and ethical pitfalls in leveraging user-generated book reviews; and, (2) professional and scholarly references that might serve as useful guidelines to avoid or manage these pitfalls.
Keywords: ·Book reviews ·Digital humanities ·User-generated content ·Social computing ·Responsible data science

Rights Metadata and Usability of Digital Collections

Trevor Stratton,

This poster provides an overview of the metadata audit that members of the Acquisitions and Cataloging Services unit (ACS) at the University Library of the University of Illinois Urbana-Champaign performed on the institution's digital collections. This audit highlighted gaps that were present within the rights metadata for items in the university's digital library, which meant that the availability of these materials for use was unclear. Through this audit of rights metadata for digitized and born-digital collections, ACS was able to identify the collections that lacked sufficient levels of description of rights metadata and worked with content creators to add this information to the digital collections.The results of this audit emphasized the fact that the understanding of what level of description is appropriate for digital library collections is not always clear. With this in mind, this poster was intended to raise awareness of the need for robust metadata in facilitating access and discovery for digital content. The lack of rights metadata, for example, results in unclear copyright statuses for the library's digital collections, which can stand as a barrier to user access as well as the findability and usability of these collections. In response, ACS continues to work toward using the results of this audit to promote enhanced description practices throughout the institution. This poster features specific data from this audit to raise awareness of similar problems of under-description and the barriers it poses throughout digital libraries and the information professions at large.
Keywords: Metadata, rights statements, copyright, quality control, metadata standards

Technology and Law in Sports Market Control

Owen Monroe, Dr Melissa Ocepek, Dr Madelyn Sanfilippo

This presentation discusses how professional sports associations' exemptions from anti-trust law control and regulate spectator access to sports broadcast markets and athlete access to team contract markets. As information scientists, we are interested in how today's landscape of sports media broadcasting and athlete labor contracts was formed through historical developments in media technology, law, and sports league policy. We conducted interdisciplinary research to find interactions and patterns of change across these domains over time. As new media technologies emerged, sports leagues expanded power over athlete labor and broadcast markets, permitted via legal rulings granting broadcast property rights and anti-trust exemptions. Building audiences and profits, leagues retained long-term player contracts, absorbed competitors, and controlled broadcasts through exclusive territories and

game blackouts. Athlete collective action later successfully challenged anti-trust exemptions, leading to free agency regulations. In the digital era, fans and policymakers have challenged blackout policies, ending certain restrictions while others persist. This presentation focuses on the history of Major League Baseball as an exemplary case in the interaction of technology, law, and sports league policy over time. I will discuss early practices in baseball employment contracts, anti-trust exemptions rulings from Federal Baseball Club v. National League and the Sports Broadcasting Act of 1961, collective action through the Major League Baseball Players Association and Flood v. Kuhn, and recent fan advocacy for open broadcast markets. A timeline of change in technology, law, and league policy will be incorporated into the presentation. This presentation comes from a paper recently submitted to iConference 2024. Our findings identify different forms of sports broadcasting blackouts and present a pattern of technological change followed by organizational action, legal support, and then response from athletes and fans. Our research provides historical precedent to current questions of how anti-trust law can be applied to emergent technology and media.
Keywords: American Sports, Anti-trust Law, Broadcasting Policy, Technology and Media History, Labor History

Three-Dimensional Archiving of Native American Artifacts at the Spurlock Museum

David Eby

In 2022, the Registrations and Handling team at the University of Illinois' Spurlock Museum of World Cultures began a project to archive Native American artifacts at the museum in a three-dimensional format. The team is currently creating 3D renders of holdings owned or on loan at the museum placed within the newly opened exhibit, Welcome to the Powwow: An Intertribal Experience. This poster presentation provides insight on processes and software trialed to create renders of fragile artifacts and objects, an explanation of how the team approached artifact handling and care of objects with three-dimensional technology, and the importance of three-dimensional photogrammetry to bring previously inaccessible artifacts to more patrons and researchers, especially given the cultural significance of Native American artifacts.
Keywords: Archives, photogrammetry, Native American

TORCHLITE: Tools for Open Research and Computation with HathiTrust: Leveraging Intelligent Text Extraction

Glen Layne-Worthey, Ryan Dubnicek, Boris Capitanu, Deren Kudeki, Stephen Downie

This poster describes a major new effort by the HathiTrust Research Center to make its rich "Extracted Features" data more accessible, more usable, and more readily available to researchers and librarians: Tools for Open Research and Computation with HathiTrust: Leveraging Intelligent Text Extraction (TORCHLITE), generously funded by the U.S. National

Endowment for the Humanities in 2021, is intended to enhance HTRC's current data delivery infrastructure with a new, non-relational ("NoSQL") database customized to host the EF dataset, and a robust and well-documented API for accessing the data within it. The API allows for retrieval of highly targeted subsets of the EF dataset, down to the level of individual volumes or even pages, if desired.This API is intended to feed lightweight analytical tools ("widgets") that can be incorporated into any modern website — for example, that of a library catalog or digital scholarly publication. These widgets can also be combined into a "dashboard," together data selection and cleaning tools, statistical information about the workset that is represented in the various widgets (e.g., total word count, most frequent words, vocabulary density), word frequency visualizations (e.g., word clouds, charts comparing frequencies of user-selected words across volumes), and comparison of multiple volumes in that workset. The dashboard and widgets are built with modern, open data visualization libraries (e.g., chart.js, D3.js, etc.).Since an important goal of the project is to allow for community-conceived, community-created, and even externally-hosted widgets, we will host a workshop and hackathon in May 2024 to attract scholars and developers from the HathiTrust user community interested in experimenting with the greatly improved access to the rich open data that HTRC has made available from the HathiTrust Digital Library.
Keywords: Digital Humanities, Digital Libraries, Community Engagement, Data

Toward Co-designed Youth Research Agendas

Rachel Magee

While more researchers are connecting with teens in participatory and collaborative ways to root research in youth perspectives and priorities, this work often focuses on specific problems or goals closely aligned with researchers' expertise areas. There are open opportunities for youth to inform research topics and trajectories from the beginning. This poster presents work conducted with 23 teens in spring 2023 that connected youth with early research experiences and collaboratively generated research ideas. Through a survey, teens participated in individual reflection on topics important to them as well as youth technology use practices, contributing a question intended for their peers on their topic and another on teen technology use. Next, participants joined in for a collaborative focus group. Participants selected a small group to refine and practice questions they were most interested in addressing with the full group. Each small group then facilitated larger discussion about their topic. We concluded the session with a full group conversation where participants reflected on what they learned from their peers, beginning preliminary analysis. I present the teen-generated topics by theme and share patterns in the ways teens sought to investigate the technology use practices of their peers. Teen-generated topics centered on three areas: Personal Growth, Obstacles, and Future Directions; Leisure, Happiness, and Artistic Pursuits; and Navigating Goals and Concerns for Society. Technology use questions centered on three topics: Assessing and Adjusting Use; Emotional Impacts of Technology Use; and Digital Community and Opportunities. I highlight how adolescents approach understanding their peers' experiences and emphasize ideas that can inform research

questions, designs, and objectives. I conclude by discussing the implications and opportunities of incorporating youth perspectives into generative discussions of future research, arguing for expanded ways for youth to inform the development of research on technology and media as well as youth culture more generally.
Keywords: teens, research topics, collaborative research

Unified Framework for Evaluating Confidence in Research Synthesis

Hannah Smith

In any field of study, existing evidence must be reviewed in order to understand the scientific consensuswithin the field and to produce actionable, science-informed goals. However, challenges in the reliabilityand synthesis of scientific evidence have been brought to the fore recently, particularly for supportingevidence-based decision making in areas such as conservation, energy policy, healthcare, andsustainable development. Bias is a challenge for all types of review as even the best review methods arelimited by their data, primary literature, and the existing evidence. Bias stemming from these limitationsmay manifest as publication bias (in which studies with certain results are favored for publishing),selection bias (in which a non-representative sample of evidence is selected for review), sponsor bias (inwhich industry-sponsored studies tend to erroneously produce results which are favorable to thesponsor), over-attention to certain topics or geographic regions, the underproduction of knowledge thatcrosses disciplinary borders, honorary authorship, and the existence of ghostwriters (which may obscurefinancial interests). While methods to evaluate the extent of this bias – in other words, methods toevaluate confidence in evidence synthesis – exist, they are difficult to carry out at scale. This project aimsto create a unified framework which can evaluate the confidence of research synthesis in a variety oftypes of review including literature reviews, systematic review, personal collections of papers, or resultsfrom subscription databases. The framework includes tools for identifying and visualizing networks ofauthorship as well as networks of citation, co-citation, and reference; tools for producing a variety ofstatistics on the bibliographies of papers (such as the distribution of publishing year and publishinglocation of references); and tools for identifying the funding sources of the literature. We envision thefinal framework to be accessible for in-field researchers, out-field researchers, and non-researchers alike.
Keywords: Evidence synthesis review confidence evaluation

Unveil: An Escape Room Web Game for Cognitive Skill Enhancement in Older Adults

Philo Wang

In an increasingly digital age where misinformation abounds, fostering the ability to discern fact from fiction is paramount. "Unveil" is a web-based escape room game tailored to address this challenge, targeting older adults. The core objective of "Unveil" is to enhance critical thinking and information evaluation skills in older adults while they navigate the intricacies

of digital information.The game employs gamification principles to provide an engaging and interactive learning experience. Players are immersed in a series of puzzles and scenarios featuring a variety of information sources, some of which may contain misleading or false content. Through gameplay, participants are encouraged to scrutinize information sources, assess their credibility, and make informed decisions.However, "Unveil" goes beyond mere entertainment. It serves as a robust research platform, capturing detailed player interaction data, including decision-making processes, response times, and problem-solving strategies. These logs serve as valuable resources for cognitive analysis, shedding light on the cognitive processes involved in information evaluation and critical thinking among older adults.
Keywords: Educational game; misinformation; older adults

Unveiling Privacy Measures in Mental Health Applications

Muhammad Hassan, Prof. Masooda Bashir

Mental health conditions have become a global public health issue, especially in the context of the COVID-19 pandemic. To cope with the increasing demand for mental health services, many people have turned to smartphone applications that offer various mental health solutions, such as therapy, counseling, and self-help. However, these applications also pose significant privacy risks for their users, as they collect and share sensitive personal and health information with third parties, often without adequate consent or transparency. In this study, we examine the privacy policies of popular mental health smartphone applications using the Fair Information Practice Principles (FIPPs), a widely recognized privacy framework. Our objective is to assess the extent to which these applications adhere to the FIPPs guidelines and to identify the gaps and challenges in their privacy practices. We hope that our findings can inform and guide policy makers and application developers to design more user-centric and robust privacy policies that ensure the safety and security of users' information.
Keywords: Human and societal aspects of security and privacy; Computing / technology policy; Privacy protections

Visualizing Race in Medicine

Chris Wiley

The Tuskegee experiment became the public symbol and face of research medical malfeasance. Numerous scholars, literature, and science frame this study through the lens of distrust, lack of ethics and informed consent. However, Tuskegee foregrounded practices, behaviors, beliefs, and standards that already existed. History and science reveal disconnections from the historical implication of black individuals and disparate experiences within communities.Currently, medical doctors cite challenges to publishing research on race within scholarly medical journals. Racial disparities within medicine, scientific records, scholarly publishing, and medical literature are not silos and are intersections. This poster is research in design illustrating their intersections and explores the following research

questions:(1) How do existing university course content describe race and ethics?(2) Can racism be conceptualized as an ethical issue within medical education?
Keywords: race; medicine, scholarly publishing; biomedical literature; racial disparities

What went wrong? Examining ethical violations in retracted papers across disciplines

Julian Chin, Professor Jana Diesner

Researchers are expected to comply with multiple regulations and ethical norms in their work. This involves adherence to institutional policies (such as Institutional Review Boards), privacy and security laws, and terms of use. However, in matters that are ethically gray but not illegal, researchers often need to rely on their own judgment, influenced by personal and professional values shaped by peers, to make ethical and responsible decisions. Prior studies highlight the challenges of making ethics-related decisions in research, given the diversity of ethics documents and discipline-specific codes as well as varied approaches to ethical decision-making in research areas. While past studies have explored the ethics codes and decision-making approaches researchers follow to make ethical decisions, there is limited literature comparing the ethical challenges faced by different academic disciplines. Therefore, this study aims to address two key questions: 1) What leads to research papers being labeled unethical and retracted? 2) How do ethical challenges related to research misconduct vary across disciplines?To answer these questions, we collected data on retracted papers featuring ethical violations from the publicly available Retraction Watch Database. Analyzing such papers offers insights into research practices that raise ethical concerns post-peer review.
Keywords: research ethics, responsible conduct of research, retraction

Who needs a main entry, anyway?

Liliana GIusti Serra, José Fernando Modesto da Silva

The action of cataloging means to describe things. Since the 19th century, with the establishment of standards, libraries have described bibliographic and authority data to allow users to identify resources by titles, authors, subjects, etc. In the traditional card catalog system, librarians faced limitations in describing data in a small and restricted space, using only one face of a standardized card measuring 5x7", with scarce space for description. This situation led to describing only essential data, avoiding a full description. Nowadays the catalogs are not constructed with cards anymore. They are on the Web, as the online public access catalog (OPAC). The tools evolved, but our describing concepts are still the same as used in the 19th century, with the 5x7" card in our heads, following patterns that no longer meet user needs. Also, over the years, the terminology of the main entry evolved as access points, but the idea of highlighting one author to the detriment of others persists, positioning added entries apart from the main entry in the OPAC visualization. This paper discusses the usage of the main entry in the library catalog. A brief history of the concept of

the main entry is made, pointing out meaningful discussions through the years. This study invites us to reflect on the maintenance of traditional cataloging practices after the advent of the Library Reference Model (LRM) and the Resource Description and Access (RDA). Also, it suggests an option for describing the access points in bibliographic records, aligning the description to LRM and RDA, and facilitating user tasks. This study is a qualitative and descriptive research based on papers, books, reports, and conferences discussing the usage of the main entry in the cataloging practice.
Keywords: Cataloging, Main entry, Library Reference Model (LRM), Resource Description Access (RDA), User tasks