Full Proposal NARRATIVE: Request RE-248850-OLS for LB21-FY21 from NC Central University School of Library and Information Sciences (SLIS)

1. Statement of National Need

Many educators and professionals in Library and Information Science (LIS) are unfamiliar with the technology and publication format for Mobilizing Computable Biomedical Knowledge (MCBK). MCBK publications include, in addition to text and graphics, "dynamic knowledge" with encodable prediction models and computable code. Such computable knowledge can provide readers (or users), for example, a diagnosis or a tool to compute a risk score for coronavirus infection from data (Friedman & Flynn, 2019). Learning how to support publications and collections with computable knowledge is the training goal of this proposal.

In general, MCBK has become an important aspect of establishing Learning Health Systems (LHSs) in the US. The vision of LHS grew out of the imperative to improve healthcare in the US and internationally where system-level challenges include the underutilization of appropriate care and the overutilization of inappropriate care despite rising costs, patient safety issues and evidence of health disparities (Friedman, et al., 2017). With training in MCBK collections and technology, LIS professionals can be contributors to LHSs and help design more effective data archives and repositories to improve information accessibility for healthcare professionals, patients, and researchers. From a pilot class to a sustainable open educational resource (OER) with a community of practice (CoP) that can use online materials, this proposal will improve delivery of computable applications to LIS professionals and healthcare providers.

The need to speed up access to information and computable tools is strong in healthcare and other sciences. Whereas printed publications provide human-readable information, new technologies and platforms support the electronic publication of computable knowledge. The availability of computable information can reduce the gap between research and practice from years to a few months. Also, data and knowledge can be part of a continued cycle in which a health problem of interest, such as Covid-19, provides data that can become knowledge (D2K), knowledge can change medical performance (K2P), and new performance can provide more data (P2D) for a continuous, true learning cycle. In the world of libraries and information centers in health clinics, medical schools, and healthcare businesses, data and knowledge could be shared on a mass action level through the application of computable knowledge objects, digital libraries and new interfaces for information that is interactive and dynamic (Friedman & Flynn, 2019).

As summarized after the 2019 and 2020 MCBK conferences, "most organizations do not yet have the infrastructure required to consume and apply computable knowledge, and

national policies and standards adoption are not sufficient to ensure that it is discoverable and used safely and fairly, nor is there widespread experience in the process of knowledge implementation as clinical decision support" (Richesson, 2020; Williams, et al., 2020). LIS professionals can contribute to defining the tools, infrastructure, and policies needed. At the 2020 conference, a four-person panel provided descriptive examples of how libraries can lead the future mobilization of CBK as co-moderated by Christopher Shaffer and Terrie Wheeler (Williams, et al., 2020). That is, although some MCBK information is already being published online and shared under creative commons licenses for repositories, new standards, policies, and sustainable metadata tools need to be defined and developed. Professional organizations and consortiums, such as Healthcare Information and Management Systems (HIMSS), the American Medical Informatics Association (AMIA), and the Clinical Pharmacogenetics Implementation Consortium (CPIC), are moving from human-readable formats to including computable, machine readable objects in their conference papers and publications. Articles applying MCBK principles, for example, may provide computable information for applying large-scale data analytics.

Specifically, the observational health data sciences and informatics (OHDSI) program supports a multi-stakeholder and interdisciplinary collaborative to create open-source knowledge based on large-scale analytics. Applying data standards for systematic research, OHDSI is already connecting almost 3000 users in 18 countries with a network of databases as presented to MCBK and LHS groups (Banda, et al, 2017). Other examples of MCBK oriented research are Dr. Grace Peng's knowledge treatment models and Dr. Herbert Sauro's call for repositories to preserve and to manage long-term digital, biomedical models. In addition, Bart Ragon has called on library science to move knowledge forward so that libraries of the future would be the gatekeeper, creator, and consumer of CBK (Williams, et al., 2020). Recent research on rapidly translating clinical guidelines for Covid-19 describes the promising use of AI and knowledge engineering to create decision making models (Fox, et al., 2020). Thus, MCBK publications are growing, and the need for trained LIS professionals is increasing.

This proposal supports the IMLS project category of a National Digital Platform by focusing on interactive, electronic publications. With MCBK experience, LIS professionals will be able to create new theories and practices for digital infrastructures of the 21st century. The National Library of Medicine (NLM) is already working with MCBK experts from the Michigan School of Medicine and Duke University. However, few LIS professionals and graduate students are familiar with the protocols and models for storing biomedical research software and data. The training will include demos (demonstrations), exercises, and hands-on labs with protected electronic health records and medical data for LIS students to learn computer and database processes for storage

and access using meta-data and reference models, phenotypes for machine learning, and how to maintain healthcare data repositories.

Diversity and Inclusion: As a prominent HBCU school, NCCU's SLIS will be able to recruit librarians and information professionals among recent and interested MLS and MIS students who represent diverse ethnical backgrounds. Health disparities are prevalent in economically challenged communities of color. Ethnically diverse students in the pilot class and future CoP will be able to support research into the application of computer programs and biomedical informatics to reduce disparities. The proposed training and resources will support the inclusion of diverse patients, providers, and researchers using decision support systems, machine learning, and AI to improve diagnoses and chronic care management. Considering effective practices for recruiting underrepresented minorities into medicine, the MCBK pilot class attendees will be provided a stipend and a social media structure for communicating and sharing as a cohort and CoP, will complete community outreach projects and activities, and will have mentors from the collaborative partners for the grant (Figueroa, 2014). The grant funding will have impact on urban and rural communities with health disparities.

2. Project Design and Timeline

2.1 **Goals and Outcomes**:

By building an online training curriculum with health informatics experts from around the country, this pilot and project will capitalize on the wealth of knowledge available within the LIS academic world. The grant would support the advancement of theory and practice in biomedical informatics while developing leaders in computable knowledge storage, access standards and tools across LIS disciplines and would address specific information needs of the medical and healthcare community currently as well as data needs in the future. Equipping LIS professionals with skills and tools in software and data storage will support the learning health systems objective that every decision affecting the health of individuals and populations should be informed by the best knowledge available.

2.2 Four Development phases

The project will start on August 1, 2021, and end on December 31, 2022. The project leaders will work with health informatics specialists and librarians as partners and resources. NCCU graduate research assistant will help collect MCBK materials and data structures. Input from collaborative partners will impact pilot

class format as guest speakers will be part of the training. The pilot training will be held online in early 2022 and be broadcast remotely.

1) August 1 – October, 2021 – Develop course plan and OER materials for 40-hour pilot based on qualitative, grounded research with collaborative partners and experts in MCBK, digital libraries and health data.
    - Share drafts of pilot modules.
    - Organize bi-weekly consultations throughout planning period with content experts (via Webex).
    - Detailed review of pilot plan in September.
    - Continued participation by project director with MCBK workgroups and grant's collaborative partners.
2) October, 2021 – February, 2022 – Advertise pilot course via social media to NCCU SLIS alums, LIS schools (including University of South Carolina and UNC-CH) and research companies; recruit 15 (plus 5 for waiting list) students who will be given stipends.
    - The pilot class selection criteria for LIS graduate students and LIS professionals will be background or future aspirations in health informatics, medical libraries, electronic medical health records or computable tool knowledge as explained in a brief statement. In addition, their plans for applying MCBK skills after the training and having basic experience with online learning will be considered.
    - Hold pilot study training (in December - January as inter-semester event).
    - Provide social media-based communications during pilot to promote cohort bonding and to connect students with mentors, who will provide experience and suggestions to 2 or 3 students during the pilot class as well as professional networking and guidance after pilot. (Statements of purpose will be requested from all recommended mentors.)
    - Observe the 40-hour training and collect data from recordings, which will include qualitative data from discussions and questions asked, and student responses; and quantitative data results from assignments.
    - Additionally, data will come from pre-class and post-class questionnaires to be analyzed for comparison and learning.
    - As follow-up there will be focus group interviews with about 5 students per focus group; sample questions would include favorite assignment and why it worked for you, suggested future exercises or improvements, personal and social results from blogging, and potential real-world applications in your current or future job.
3) March, 2022 – May, 2022 –Analyze learning data from observations, measurements, follow-ups, and social media.

- Concerning the social networking, for example, a social network analysis or SNA of sharing would be developed, and a word cloud of key terms to indicate most commonly used terms would be created.
- Social analysis will also define significant professional connections and future MCBK activities from among students, mentors, and instructors that cohort indicates will be pursued after pilot class as gathered from data-sharing information from the required, limited-access blogs and focus group interviews.
- Data listing key learnings from pilot class experience will be documented. Prepare collaborative reports on pilot results and future project design from current and potential stakeholders.
- Stakeholders are knowledge developers and researchers on MCBK workgroups; LHS authors and teachers; IMLS grant collaborative partners; guest speakers during pilot; interested professional organization members of the Medical Library Association, Special Library Association, and Association for Information Science and Technology; NCCU library staff and OER supporters; and NCCU-SLIS students, faculty, and interested colleagues.

4) June, 2022 – December, 2022 - Present pilot results and future project plans.
    - Give conference and workgroup presentation to 2022 MCBK meeting in Bethesda at NLM.
    - Summarize learning project in an article for an LIS journal (such as *JELIS*), or LHS journal, present at a conference (such as ASIS&T or SLA).
    - Complete open-source education documents and development of course materials for the sustainable, online training resource with Community of Practice (CoP) support during this reporting and dissemination phase.

2.3 Overview of Training Pilot Class:

The pilot class will be online with recordings and digital materials. Based on the following Student Learning Outcomes (SLOs), the training will provide 40 hours of discussion, hands-on exercises, and special guest speakers from December 13 to January 10, an inter-semester mini-session. For example, a 2-hour introductory orientation on Monday, December 13, nine 4-hour classes Tuesday-Thursday, December 14, 15, 16, 28, 29, 30, and January 4, 5, 6. A 2-hour summary and closing would be on January 10.

The preliminary SLOs include being able at the end of the training to:

- Define MCBK;
- Summarize and critique sample MCBK articles and publications;

- List advantages and disadvantages of computable knowledge in a learning health system (LHS);
- Prepare a sample annotated bibliography from the NLM of MCBK and computable publications;
- Define and use computable (machine-executable) code from a publication;
- Develop online procedures for building and accessing an MCBK collection;
- Write a report on the Observational Health Data Science and Informatics (OHDSI) Data Collaborative and use of electronic health records (EHRs);
- Demonstrate the use of MeSH metadata of MCBK collection samples;
- List examples of clinical decision support systems that utilize MCBK;
- List professional organizations and their feedback on MCBK standards, such as AMIA, ACOG and HIMSS;
- Summarize security and risk-assessment issues with MCBK materials;
- Describe how MCBK and knowledge sharing are advancing healthcare.

To measure learning success, the project director/co-director will provide written feedback to students and records. Also, data from the pre-class and post-class questionnaire will be used to measure learning. Students will have 80% completion of work as minimal requirement to receive 4 Continuing Education Unit (CEU) credits for 40 hours of education from NCCU.

2.4 Storing for Open Access Resources from the Pilot Class:

The online pilot class recordings and digital materials will follow OER guidelines to remain resources that that are openly accessible. With assistance from the Assistant Director of Library Services at NC Central University (NCCU), Ms. Jamillah Scott-Branch, MCBK training resources will be collected, stored, and made accessible. NCCU associate professor and OER consultant, Dr. Patrick Roughen, will review creative commons licensing and agreements. Following the “5Rs” of OER, the MCBK materials will be retained, reused, revised (as needed), remixed (as needed) and redistributed according to approved processes.

3. National Impact

The MCBK training project will be timely with strong impact during the 2021-2022 continued development period for computable biomedical knowledge and standards for information collections, metadata, models, and interactive tools. For many years the Library and Information Science (LIS) world has become intimately involved with the information needs of medical and healthcare professionals. Health Informatics programs have been implemented in a variety of LIS schools, such as CHIP at UNC-Chapel Hill. However, research has found, "little attention has been paid to forecasting the information resources and services that researchers,

specifically, will need" (Cain et al., 2016; doi:https://doi.org/10.5860/crl.77.5.595). Whereas, LIS has begun to lose reference and knowledge management positions to more IT-based workers, this grant would address the need directly by involving LIS professional on the ground floor as MCBK repositories, protocols for software and data, and AI/Machine learning models are being designed and implemented for rapid implementation. LIS professionals in medical libraries and public libraries where users search of healthcare information will be able to keep a competitive edge within the data archives and repository management field for biomedical and other sciences that use computable information.

By building relationships and skills early, the LIS field can impact information policy and digital access decisions as they are being made rather than having to react to poorly implemented policies or metadata structures put in place by decision makers who are more focused on the technology than on information itself. As NLM director, Dr. Patricia Flatley Bennan has said, we need models with rigor, reproducibility and reuse (Richesson, 2020). Federally funded data research collections are accessible, but data and programming need to be shared. The standard approach in US libraries towards nomenclature, format, and transmission is not sufficient. LIS needs to prepare to lead in new directions that support community healthcare and collaboration to archive and store software and data as well as printed, text-based and graphical objects. The code and data summarized in research publications needs to be accessible for validation of software and data analytics. The IMLS can support leadership who will improve ease-of-use for research data in biomedical research and other fields.

Most importantly, promoting healthcare research in libraries has been a national concern for many years. A Health Information Initiative from the National Network of Libraries of Medicine (NNLM) stated in 2018 (http://www.ala.org/pla/initiatives/healthliteracy) "access to current and reliable health information is imperative for the well-being of all Americans …." An LHS depends on collaboration and trust if healthcare data are to be shared. Thus, learning how to make data that is transparent in format (or phenotype) with formulas or algorithms that can be repeated will allow LIS professionals to support an effective learning health system because research data will be reliable and verifiable. MCBK technology and principles support skills that ensure databases are interoperable with metadata and API access and follow standards to support the preservation of information and data. The intent of this proposal is to provide a pilot course for approximately 15 LIS professionals, and to develop sustainable, online training about MCBK and health data. The IMLS grant will help meet national

healthcare needs, expand education goals for library and archive professionals, and provide access to diverse digital content about health research.

Finally, as the impact of the Covid-19 pandemic has shown, the importance of rapid dissemination of data, models, and computable information can be crucial during a crisis. Early research during the pandemic demonstrated the use of AI and knowledge engineering to create decision making models (Fox, et al., 2020). Furthermore, the need for an accurate and accessible public health data infrastructure during such a crisis underscores the absence of data sharing; MCBK and LHS researchers are publishing frameworks and data analysis results for transmission dynamics with early models and metadata for data sharing agreements and standards (Foraker, et al., 2020). There is a strong national need for LIS professionals to contribute to the development and implementation of MCBK publications, libraries, data repositories, and computable information tools that are user centered. This proposal will provide training and sustainable learning resources developed and tested in a pilot course and presented for validation in scholarly journals and conferences.

4. Resources: Personnel, Partners

4.1 Key Personnel

Dr. Deborah Swain (Project Director) contributes research and expertise in knowledge management (KM), data analytics and health informatics; she serves on MCBK workgroups for trust and sustainability. She has her PhD. And is an Associate Professor in the School of Library and Information Sciences at North Carolina Central University in Durham, NC. Dr. Swain has over 20 years of experience in business, education, and managing information projects. She has also taught at UNC-Chapel Hill, NC State, and Campbell University. She does research in knowledge management, health informatics (telehealth) and usability (user experience). She completed her doctorate in Information Science at the University of North Carolina at Chapel Hill (1999). She has an MA from UNC-CH in English and a BA from Duke University. She teaches graduate courses in information science, health informatics, human factors, data mining, knowledge management and artificial intelligence (expert systems). She has worked as a systems engineer, information developer, consultant, and project manager for corporations, such as IBM, AT&T, and Bell Labs. Dr. Swain is an active member of the American Society for Information Science and Technology (ASIS&T), International Conference on Knowledge Management (ICKM), American Library Association (ALA), North Carolina Library Association (NCLA), and

Special Library Association (SLA). She was Conference Chair of ICKM 2020 in December.

Dr. Chris Cunningham (Co-Project Director) has CIO and program director experience in business and academics; he teaches graduate MLS courses in management and leadership. He is an Adjunct Professor in the School of Library and Information Sciences at North Carolina Central University in Durham, NC and a Data Analyst for Cutting Edge Signs in Concord, NC. Dr. Cunningham has almost 25 years of experience in business, education, and managing information projects. He has taught at UNC-Greensboro, Kent State University, and the University of Tennessee. He does research in digital divide issues as well as working on several GIS papers. He completed his doctorate in Library and Information Science at the University of South Carolina, his MLIS from UNC Greensboro and a BA from UNC Charlotte with minors in Mathematics and Physics. He has worked as for over a decade in the business world, building several information projects to increase productivity while reducing costs. In addition, Dr. Cunningham worked as Program Coordinator at UNC Charlotte, working on recruiting, and retaining students for the Graduate Business Program Office. He is an active member of the American Society for Information Science and Technology (ASIS&T), with past accomplishments including chairing SIG -ED and an officer for SIG DL. He is a former member of American Library Association (ALA), North Carolina Library Association (NCLA), and Special Library Association (SLA). Note: His background in education, administration, and coordinating student travel grants will be valuable in planning and implementation of the MCBK training.

4.2 **Collaborative Partners:**

Dr. Charles Friedman, LHS department chair, and Rachel Richesson, Michigan School of Medicine; Kathleen Young, Editorial Assistant, *LHS* journal; health informatics, LIS, and data analytics consortium partners in NC (Duke, UNC-Chapel Hill, East Carolina University, UNC-Charlotte and Wake Forest) including Javed Mostafa at UNC-Chapel Hill; LIS, KM and HI experts at Arizona (Jerry Perry), the University of South Carolina (Feili Tu-Keefner) and North Texas University (Suliman Hawamdeh, Ana Cleveland and Jody Philbrick); NIH data experts consulting at the NLM: Dawaei Lin; Robin Ann Yurk at MD Yurk; Director of Libraries for the Medical University of South Carolina, Shannon Jones.

Note: Short biographies of the primary collaborative partners are provided in the List of Key Project Staff and Consultants.