A Librarian’s Guide to The National Center for Biotechnology Information

By Tina Griffin

The National Center for Biotechnology Information (NCBI) is a division of the National Library of Medicine (NLM) and the National Institutes of Health (NIH). It has the mission to "develop new information technologies to aid in the understanding of fundamental molecular and genetic processes that control health and disease.”1 Established in 1988, NCBI focuses on research into bioinformatics and computational biology, engages the scientific community for education and training, and develops/supports databases and software in these fields, among other activities. NCBI resources topically span the range from nucleic acids, genes, and genomes to protein, taxonomy, chemicals and bioassays. All of these are accessed online through their website at www.ncbi.nlm.nih.gov/.

NCBI has many outlets for education and training on their resources. They offer webinars, courses, and tutorials in various formats; present at conferences; and provide thorough documentation through handouts, fact sheets, and FAQs. As part of their outreach in education and training, NCBI understands that librarians are heavily involved in the dissemination of biomedical literature through one of its databases, PubMed. However, it also recognizes that librarians have a larger role to play in searching for biomedical data itself and in the education and training of researchers, students, and the public in regards to NCBI’s other databases and resources. To equip librarians to be able to educate and promote NCBI resources, they offer the course A Librarian’s Guide to NCBI each year. This course is specifically designed to support librarians as they begin to implement bioinformatics services into their service model.

Admission requires an application, statement of intention regarding how this education would support your library’s goals, statement of support from a supervisor, and a recent copy of your CV2. Acceptance is limited to between 25-30 people, and includes librarians with no experience in any science field, to those that have advanced degrees in a STEM discipline. Those who are accepted into this program have their course costs covered. However, they are required to sustain themselves while attending the in-person sessions. Food, travel, lodging, and incidentals are their responsibility. Participants are required to bring a laptop.

A Librarian’s Guide to NCBI is a MLA continuing education credit course taught in two parts. The first part is Fundamentals of Bioinformatics and Searching. The goal of this section is to give an introduction to bioinformatics and molecular biology concepts and an introduction to NCBI tools. This part is taught asynchronously and online by Diane Rein, librarian and scientist from the University of Buffalo. The information covered includes:
• The origin of bioinformatics and bioinformatics databases
• Bioinformatics as a research discipline
• An NCBI overview
• A brief look at some of the bioinformatics records within NCBI and exploration of the relationships between databases

Content is delivered through Moodle, an online content management system designed for online classes. The instruction consists of readings, and recordings of lectures and other keynotes talks. Class assignments are vocabulary and exploration exercises. The vocabulary exercises report on dictionary and reference definitions of molecular biology terms. The feedback on this exercise is particularly helpful, as the answer key was an aggregate of answers from the whole cohort as many people took different approaches to the question. The exploration exercises are guided walkthrough tasks that reported out answers to those specific tasks. The feedback on these exercises is that the exercise would not be considered complete until all answers were correct. Participants are given feedback and allowed to resubmit continually until the due date. Completion of this portion of the course is worth 18 MLA CE contact hours.

The second part of this course is the formal Librarian’s Guide to NCBI, held in-person at the National Institutes of Health conference center in Bethesda Maryland. This part is five full days of lecture and exercises similar to the first part. The information covered includes:
• Biology overview
• Introduction to the Entrez bioinformatics inquiry system
• Sequences, assemblies and genome overview
• BLAST searching
• Sequence variation and searching
• Gene expression and Biological pathways
• Bio molecular structure and searching
• Drug and small molecule searching
• PubMed updates
• Public access compliance, MyNCBI, and SciENcv overviews

Content is taught by individual instructors from NCBI and Diane Rein. In addition to instruction, the teachers are developers and administrators for the NCBI databases and tools. The five day course is broken down topically. Each day has a morning and afternoon session taught by one or more instructors. Typically, instruction consists of lecture with intermittent hands on guided exploration or problems sets. Completion of this portion of the course is worth 36 MLA CE credits.

One of the top advantages to taking this course is to get the training first hand from the researchers and instructors at NCBI. The suite of NCBI resources is a complicated, interconnected, massive amount of bioinformatics information that is not intuitive to navigate. In addition, the individual records for the databases can require a certain amount of specialized knowledge or expertise. The instruction sessions for Librarian’s Guide to NCBI are grouped according to topic and taught by the NCBI personnel that works with them. All of the instructors have intimate knowledge of the NCBI suite of databases and
tools, and all of them specialize according to their background and training. In addition, the course offers support and training beyond the in-person class. Once participants have “graduated” from the course, they are moved to a graduate Moodle. There, participants communicate through message boards to their cohort and to the previous years’ participants. In addition, communication is regular through the Moodle. Many participants contribute reference questions to be crowdsourced or ask general questions. Lastly, the NCBI instructors hold monthly office hours for graduates to continue training on resources. This time is also used by previous participants to demonstrate how a new resource can be used with NCBI tools or data, or to demonstrate the instruction they have built for their institutions.

Related to content is the ability to get a minimal biology education, or refresher. Even to those who have education and experience in molecular biology, bioinformatics associated disciplines change rapidly. Many researchers can keep up with the subject when it intersects with their work; however, it is more difficult to keep up with the discipline as a whole. Having dedicated time to focus on refreshing or gaining biology knowledge will help librarians have effective conversations with their faculty. Having that current biology knowledge will also inform them of what may be needed regarding implementing bioinformatics services in terms of tools and support.

Another advantage to this course is the ability to network with other science librarians working within the bioinformatics sphere. I attended the NCBI course asynchronously in January and in-person in March of 2015. There are many people that I still keep in contact with, have seen at conferences, and have exchanged instruction materials with because of this course. As a new librarian and liaison to graduate studies in the health sciences, this has been invaluable.

Despite all the benefits to the course, there are drawbacks. The first being the time commitment necessary to complete it. The course requirements are firm on assignment deadlines and participation, particularly for the asynchronous portion. Those that missed a deadline were locked out of further course content and further prohibited from attending the in-person course. Keeping up with the schedule in either the synchronous or asynchronous portions require putting this course as a priority above other tasks. If work hours are not allocated to complete the course, potential participants need to carefully consider their capacity and willingness to do the work “on their own time.”

Another concern is the issue of not having a science background. Having a science degree wasn’t a requirement for the course. Some of the topics and biology concepts are difficult to conceptualize and understand without that background. The instructors work with these tools and databases full time and have years of institutional knowledge working with them. There is no way to cover enough content to make a person feel competent in any topic, in a single half day session, as in the case of the in-person component. That said, some sessions could have made learning easier by updated teaching methodologies rather than from the lecture format.
Lastly, the site location was a concern. It was a great experience to be at NLM and NIH. There is time enough in a day to explore the campus and the library, however, the sessions are held in the conference center, where the power and wireless capability are sorely inadequate to perform rigorous work. There was not enough power (outlets) or internet capacity for all devices to run simultaneously. Frequently, instructors ran the search before the participants so that the search would not time out. And the connectivity of the wireless was inconsistent and required frequent permissions/reconnections. To my knowledge, this is being addressed in future offerings.

Overall, *A Librarian’s Guide to NCBI* is a valuable professional development experience for libraries and librarian’s looking to incorporate bioinformatics into their services. Depending on the level of experience with the subject matter, a librarian taking this course will be able to enhance service offering that include workshops on through research services.


Tina Griffin
Visiting Assistant Professor and Information Services and Liaison Librarian
Library of the Health Sciences-Chicago
tmcg@uic.edu