Not your usual content: the Rosarium Project. An online collection of materials on the Rose using the Text Encoding Initiative and oXygen.

Julia R Tryon, Providence College
Not your usual content: the Rosarium Project. An online collection of materials on the Rose using the Text Encoding Initiative and <oXygen/>.

Academic librarians with faculty status, such as myself, typically must have a research agenda to enable them to achieve tenure and promotion. Many, though not all, focus their research on topics in library and information science. My chosen specialist subject however is roses.

I am captivated by roses as plants, as things of great beauty, and as objects of decoration and adornment. I am interested in rose garden design. I want to know all those rose gardening tips, tricks and techniques. And I find those insanely devoted rosarians fascinating. But I am most interested in what has been written about roses.

Just so you don’t think me crazy in my love for roses, consider that 4,000 year old tablets from Sumerian royal tombs list rose preparations such as rose water and attar of roses side by side with servants and food as items necessary for the afterlife. Emperor Nero liked to have rose petals showered down upon his dinner guests and reportedly spent something like a quarter of a million dollars strewing rose petals along a beach. And Empress Josephine amassed a world-renown collection of roses at Malmaison during the Napoleonic Wars with help from both the French and the British. Napoleon ordered his men to search all seized vessels for roses and send them to Josephine. The British helped out by allowing rose plants and seeds through their own naval blockade.

In 1849, the Thomas Rivers Nursery in Sawbridgeworth, England alone sold 8,000 bushes of this rose ‘Géant des Batailles,’ and that was only one of the hundreds of varieties available at the time. English rose fanciers founded the Royal National Rose Society in 1876 and Americans followed suit with the American Rose Society in 1892. So I am not alone in my fondness for roses.

A lot has been written about roses over the centuries. Information on roses appeared in herbals of the sixteenth century. Articles on roses appeared in the 1700s and the first book in English solely about roses was printed in 1796. Writing on roses really took off in the early nineteenth century. There were hundreds of books in numerous editions and innumerable articles about roses published throughout the century.

At this time there were several things happening that account for this upswing. First, there was a boom in exploration including expeditions by famous plant hunters such as Lewis and Clark in the American West and Robert Fortune in China who sent back specimens of new roses to nurserymen and their wealthy clients for breeding and propagation.
Secondly, there was the Industrial Revolution and with it a burgeoning middle class with time and money to spend, and a new suburbia in which to spend it. The new middle class with its newly acquired suburban plots suddenly took to gardening. Not the mere growing of fruits and vegetables but decorative gardening which had previously only been an option for the rich and famous.

In any proper 19th century garden, there were, of course, roses. David Stuart, in his book ‘The Garden Triumphant,’ explains “…the queen of the garden was the rose; indeed, roses are almost the Victorian flower…Every garden that could…had roses, from the humblest cottages to the very grandest mansions.

Frank J. Scott, who wrote 1870’s popular book The Art of Beautifying Suburban Home Grounds, would have agreed for he wrote:

“We have not previously mentioned the Rose, among flowers and bedding plants, for the reason that, being the queen of flowers, more than ordinary attention is usually considered due to her. Besides, her royal family are so numerous, so varied and interesting in their characters, and have been the subject of so many compliments from poets, and biographical notices from pens of distinguished horticulturists, that it would be presumptive to attempt to describe, in a few brief paragraphs, the peculiar beauties and characteristics of the family; still less of all its thousand members.”

Lastly, advances in printing, including chromolithography and the decreasing costs of paper and postage afforded the dramatic increase in the publication of books, magazines and catalogs. Nurserymen were able to put together huge catalogs of new plants with colored images and send them and their plants all across America. They took to publishing articles on gardening in their catalogs and then began issuing gardening magazines. Nursery owners in America wrote key rose books such as Robert Buist and his ‘Rose Manual’ of 1844.

So what’s a librarian with a love of roses supposed to do? Back in 1992 when I first started my project on the genus Rosa, the typical way for a librarian to share her love and interest in something was to write a bibliography, annotated or otherwise, of books and articles. So I set to work and collected hundreds of citations for books and articles on roses. But then other projects got in the way and work came to a halt.

In early 2015, I decided to go back to working on my rose project. However things have changed and the end product will no longer be a printed bibliography, since they are a thing of the past. Instead, it will be a fully searchable database of full-text materials with images, a glossary, and of course a bibliography, on the Web. The project is called the Rosarium Project.
It will include books, articles from popular magazines, items in scholarly journals, columns from newspapers and gazettes, pamphlets, and catalogs. Because of the full-text nature of the database and copyright considerations, my focus is on materials appearing before 1923. I have also chosen to limit the database to materials written in English and to items of non-fiction. I have thus far identified over 700 documents and have the citations for them safely stored in a bibliographic utility.

You may wonder why am I restarting my project now. Well, for starters, the projects that got in the way of research back then have finally wound up. Also, technological changes since the 1990s make it possible for a mildly tech savvy person to do great things and make them available to the world.

My background is in database searching and the thought of creating a searchable database is exciting. Since I have an interest in what has been written about the rose, I feel that others too would appreciate being able to access information on roses that was used in the past. In particular, I feel there is a need to know about roses from the past that were hardy without modern chemicals, about gardening techniques that have been forgotten but which may be greener than rose growing practices nowadays and about historic garden design.

The intended audience for the Rosarium Project includes rose enthusiasts like me who are interested in all things roses. But also ordinary gardeners interested in adding roses to their floral palate, and scholars, particularly those interested in the history of gardening and horticultural science, and scholars of popular culture focusing on leisure studies and suburban life.

I’ve begun the project with low hanging fruit. I have started by searching the Reader’s Guide to Periodical Literature. Since the Reader’s Guide uses standardized subject headings and my library has access to the Reader’s Guide Retrospective database from Ebsco, research was simple. The designated subject heading was simply “roses.”

I had to review each record found to remove records for poetry and other literary genres. There were 163 articles of non-fiction indexed for the period from 1894 to 1922. I have collected all 163 articles through interlibrary loan and other means.

31% of the citations collected were from one journal. Not surprisingly, it was The Garden Magazine which was published between 1905 and 1924. The magazine had twenty regular departments and one was for roses.

Twenty six articles appeared in magazines dealing with country life. The most prominent being Country Life in America. Another 22 articles came from journals on domestic architecture and
landscape design. A surprising number appeared in general interest periodicals, literary reviews, and art journals. The subject matter runs the gamut from new rose introductions, to basic gardening techniques, to rose petal recipes.

In order to realize my vision of turning these articles into a full text online repository, I have decided to transcribe the myriad texts I uncover and encode them using the language developed by the Text Encoding Initiative Consortium and expressed in their P5 guidelines, otherwise known as the TEI. Or in other words I am using the TEI to markup the full text of my resources in an XML file so that they can be collated into a fully searchable database available on the Web, using Extensible Stylesheet Language Transformations or XSLT. Providing full text searchability permits the widest use of the documents by researchers.

TEI is a common language in the Digital Humanities for scholarly markup of texts which allows for intricate textual analysis. Marked up texts are machine-readable and manipulable which allows scholars to compare and search large bodies of texts. TEI uses XML which is easy to learn and use. It is adaptable for different types of materials, such as prose, poetry, plays, manuscripts, and letters. It permits standardization of spelling while still providing access to the original material. Also encoding texts with TEI allows me to add value to the transcripts by pulling information in from a bibliography, a personography, and a glossary. I can also insert images that accompanied the text as well as append images of the original work. In addition I can link out to relevant materials on the Web.

A TEI encoded file must contain a header and a body. The header contains the metadata for the electronic document such as the title, information on the original source of the text and its electronic publication, as well as the people responsible for the work. The header can include a project description and encoding decisions as well as revision statements and the taxonomy used. The body is where the transcription of the written material is housed but can also include images, and front and back matter such as notes, bibliographies, glossaries, and indexes.

I initially learned the TEI by studying and following examples of TEI encoded files. One great resource is the TEI by Example Website (http://teibyexample.org/TBE.htm).

I use the XML editor called <oXygen/>. One could use a basic XML editor but <oXygen/> knows the language of the TEI and provides instant validation as I encode my transcriptions. It provides encoding prompts and shortcuts. The license for <oXygen/> for academic use is inexpensive.

I use the TAPAS Project to see how my encoding renders on the Web. The TAPAS Project hosts TEI projects and displays the files using the TEI Boilerplate. It is an inexpensive way for
projects to be published on the Web. Using the site requires no expertise in things such as CSS and XSLT. It also obviates the need for hardware such as servers, and specialized IT personnel.

The project’s workflow begins when I transcribe each article using a word processing application. I could transcribe directly into <oXygen/> but I am more comfortable using a word processor and like working in a cleaner workspace. I then copy the file and paste it into the body of an <oXygen/> XML file.

When I am encoding a file, I consult information on the TEI website regarding the element I wish to use. Each element has its own webpage, which provides details on where and how the element can be used and examples to follow. At the same time I also have open the original PDF so I can compare it to the transcription in <oXygen/>. I constantly toggle between the TEI website, the PDF and <oXygen/>.

In <oXygen/>, I use the TEI to tag names, places, varieties and colors found in the body portion the file. Names are linked to a TEI encoded “personography” file. A personography is to people what a bibliography is to books. At the same time, I link rose varieties to terms in an encoded glossary file. The glossary includes rose varieties, botanical and common names of roses, and other rose related topics of interest such as rose organizations and famous rose gardens.

I place a link to my encoded bibliography file in the TEI header. Each article is then assigned subjects from a taxonomy to improve findability. I have created my own taxonomy for the project as Library of Congress subject headings were insufficient. Once I have completed the text encoding process, I then upload the XML file into the TAPAS project to make sure I’ve encoded things properly. The TAPAS Project is a temporary home for the project. In the near future it will have its own web presence.

In phase two, I will be revisiting the journals indexed by the Reader’s Guide since I have discovered it was not completely rigorous in its indexing of rose articles. To do this I will rely on resources such as the Hathi Trust, Google Books, the Internet Archive, and library collections.

There are a number of resources for identifying further rose material. In addition to the aforementioned resources, I plan to also use the only two bibliographies on roses in English, Poole’s Index to periodical literature, WorldCat, materials digitized by the American Antiquarian Society, the Biodiversity Heritage Library, and special collections in libraries around the world.

As far as expenses and investments go, I’ve spent about $5,000.00 so far. The largest outlay was for a 27” iMac with retina 5K display. And it’s worth every penny. I really need a large screen
because I am having to refer to information located in a variety of applications constantly. That means a lot of windows open at once.

The license for <oXygen/> costs only $99.00, membership to the TEI Consortium which is required in order to use the TAPAS Project is $50.00 per year, and I’ve spent $300.00 for a TEI Workshop run by the folks behind the Women Writers Project at Northeastern University in Boston, MA. But the biggest investment in this is time.

I really enjoy reading the old articles and encoding them with the TEI. Which is a good thing since I foresee working on this project for the next fifteen years. I hope that you have found my project interesting and I would be happy to answer any questions.