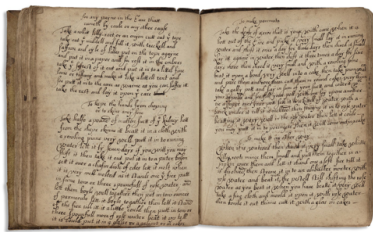


CROWDSOURCING FOR SHAKESPEARE

By Roberta Kwok

January 16, 2017



Shakespeare's World, a new project on the research crowdsourcing site Zooniverse, aims to transcribe thousands of pages of difficult-to-read manuscripts. PHOTOGRAPH COURTESY FOLGER SHAKESPEARE LIBRARY

Around 1675, a woman named Margaret Baker wrote out a remedy for aches whose active ingredient was a puppy. “Take a whelp that sucketh the fatter the better & drowne him in water till he be deade,” she advised. The reader should then gut the dog, fill its belly with black soap, “putt him one a spite & roste him well,” and apply the fat drippings to the patient’s skin, wafting the scent of warmed sage over him at the same time. “It will helpe him by the grace of god,” she concluded.

Baker’s prescription is on one of thousands of pages of handwritten documents from sixteenth- and seventeenth-century England that volunteers around the globe have transcribed as part of an initiative called [Shakespeare’s World](#). Launched on Zooniverse, a research-crowdsourcing platform, in 2015, it is an effort to better understand everyday life and language around the time of William Shakespeare. The idea arose in 2013, when Chris Lintott, Zooniverse’s founder and an astronomer at the University of Oxford, asked his friend Victoria Van Hying, a scholar of English

literature, what she considered the most pressing problem in the humanities. “I said, ‘It’s definitely text transcription,’ ” Van Hyning, who is now Zooniverse’s humanities principal investigator, told me recently. Researchers have amassed enormous collections of old handwritten documents, but lack the time and resources to transcribe them all.

The problem was familiar to Lintott. Zooniverse grew out of a site called Galaxy Zoo, which Lintott and his colleagues created in 2007 to help them classify images of galaxies by shape. “We discovered by experiment that a Ph.D. student will only look at fifty thousand galaxies before they tell you what you can do with the rest of them,” Lintott said. Galaxy Zoo allowed volunteers to help sort the images, and it was eventually expanded into Zooniverse, which other researchers have since used to identify possible planets outside the solar system and to classify animals in the Serengeti. Applied to old manuscripts, the same strategy would allow researchers to build a repository of transcriptions that could be searched for quantitative answers to historical questions—how often rosewater was used in plague medicines, say, or when chocolate began appearing regularly in recipes. Similarly, linguists could trace the evolution of English in more detail. The first-known records of many words are in Shakespeare’s plays, but it’s not always clear which he invented and which were already commonplace. The handwritten material of Shakespeare’s contemporaries is “more or less hidden,” according to Laura Wright, a historical linguist at the University of Cambridge and a Zooniverse volunteer. “Of course it looks like Shakespeare invented all this stuff, because his stuff is in print,” she said.

To tackle the problem, Zooniverse partnered with the Folger Shakespeare Library, in Washington, D.C., and with the Oxford English Dictionary. Volunteers for Shakespeare’s World can view images of documents from the Folger’s manuscript collection, including family correspondence, household recipe books, and letters by state officials, and transcribe as little or as much of a page as they want. Many people initially struggle with the strange letter shapes and abbreviations, and the unstandardized spelling. “I humbly take leave” might be spelled “i umbli tacke leue”; “yolks of eggs” might be “yealks of egs.” And some handwriting is downright messy. “It looks like a spider’s gone in the ink and crawled all over the page sometimes,” Tracey Dixon, a civil servant in York, England, told me.

Fortunately, individual users do not have to worry about getting the text perfectly right. Multiple people transcribe each line independently, and an algorithm originally designed to identify similar DNA or protein sequences compares the strings of letters to determine a likely best answer. If most users agree on most of the text, the line is considered done; otherwise, the program keeps gathering data from more people until a consensus is reached or “we cry mercy,” Lintott said. (Those difficult cases are set aside for experts.) Spot-checks suggest that the quality of completed lines is close to that of scholarly work, Van

The first-known records of many words are in Shakespeare’s plays, but it’s not always clear which he invented and which were already commonplace. PHOTOGRAPH BY STOCK MONTAGE / GETTY PHOTOGRAPH BY STOCK MONTAGE / GETTY

Hyning said.

So far, around twenty-five hundred Zooniverse users have completed more than thirty-three hundred pages. Often, once volunteers get the hang of transcribing, they stick around to learn historical tidbits, follow the authors' personal dramas, and contribute to research. "Even though you're having fun, it feels much more justified than trying to catch all the Pokémon or whatever it is people do these days," Dave Henderson, a freelance archeologist in Edinburgh, told me. At the same time, volunteering has its moments of tedium. Henderson now works primarily on letters rather than recipes, because one ingredient kept cropping up. "If I have to transcribe the word 'rosewater' one more time, I'll go mad," he said. "Some of the quantities—it must have absolutely reeked of roses."

The documents reveal the anxieties of early-modern life, many of which feel remarkably familiar, even if the solutions aren't. A remedy for scaly skin calls for rubbing a mixture of butter, mercury, and "fillth of a dogge" (a possible reference to dog feces) on the forehead. Readers are advised to suck beer out of a quill to avoid imbibing too much, to treat hair loss with oil of sulfur, and to concoct a drink for "Melancholye and weeping" by simmering rosemary flowers, sugar, and claret wine over a fire. One letter writer claims that he is too frail to visit a sick friend, explaining that he cannot "goe to horsbacke wthout the helpe of too or three att the least." "You actually get quite cross with him," Elisabeth Chaghafi, an English lecturer at the University of Tübingen, in Germany, said. "Come on, couldn't you be bothered to ride there and see your dying friend?" Dixon said that she wept while transcribing a copy of a letter from Sir Walter Raleigh to his wife, written when he thought he was about to be executed. (She was less impressed with Sir Francis Bacon, who "whinges all the time about not being given the particular jobs he wants.")

Already, the project has yielded linguistic discoveries. Volunteers have found recipes for "Taffytie" and "Taffity" tarts, which might be variations on "taffeta," implying a delicate texture. Combined with an existing record of a similar usage in the O.E.D., the new examples suggest that this was an established genre of dessert, like lemon-meringue pie is today, according to Philip Durkin, the dictionary's deputy chief editor. A volunteer came across a recipe for "portugall farts"; Durkin noted that the O.E.D. already contains the phrases "Fartes of Portingale" and "ferte of Portugall," defined as "a ball of light pastry," but "to have 'portugall farts' as well is good," he said. One letter, from 1567, about a headstrong youth uses the term "white lie," pre-dating the O.E.D.'s earliest record of the phrase by nearly two centuries.

The Zooniverse project is slated to last for at least two more years, and in the meantime the Folger team will be uploading finished transcriptions to their Early Modern Manuscripts Online database, which launched in beta this month. Volunteers are continuing to log hours in the evenings, on holidays, and during coffee breaks. Heather Wolfe, a curator of manuscripts at the Folger, said, "I'm just so touched by their devotion."

More: [Crowdsourcing](#) [William Shakespeare](#)

The stories that matter.
The humor you need



Subscribe for \$4 a month.

Subscribe

Cancel anytime.