Austin N Fife - Letter of Qualification

Dear Members of the Search Committee:

My name is Austin Fife and I believe that my experience curating data with R and the Tidyverse make me a strong candidate for the position of Data Curation Technician - SP002970P:

Experience in working with tabular data. I used R and the tidyverse packages frequently to create, curate, manipulate, analyze and validate data, during my MSc and PhD in Entomology. I often clean and managed multiple files simultaneously, for example, I had recently had to clean and combine over 50 .csv files of 'dirty data', which were imported en mass, cleaned and merged into a single data sheet for analysis. I have worked with tabular data with R and the Tidyverse packages for the execution of over six scientific research projects for six years.

Experience with spreadsheet software. I have used multiple spreadsheet software including Excel, Google Sheets, and Libre/Open Office over the last ten years to record experimental data from agricultural experiments and various other projects.

Experience with written communications such as reports, manuscripts, conference abstracts, and grant applications. I have extensive experience writing technical documents of varying lengths and formats: My research has been published on a regular basis, including 2 journal articles in peer-reviewed journals, a technical report, and a manuscript in revisions for the Journal of Integrated Pest Management. I also have created and presented conference abstracts at over six scientific conferences, and I have written a few grant applications, one of which was funded (see CV).

Experience using git for version control. I have been actively relying on git and GitHub daily for the past year for version control of my code, data, and manuscripts related to my research.

Experience in bulk import of files, merging data sets, automated error checking using command line software such as R, bash, python, or OpenRefine. I have imported large numbers of files, merged data sets and checked for errors with R and the tidyverse package for each of my major research projects for the past six years. I also am familiar with bash, and have a basic understanding of Python. I would be able and willing to pick up other programming languages as needed.

Experience in R Programming with R base commands and the R Tidyverse. I program with R daily to accomplish my research needs, including data visualizations, statistical analysis, data cleaning and merging. I rely on a combination of base R functions, and I use the R Tidyverse for every one of my research projects. This frequent use of R has taught me how to troubleshoot common problems with data cleaning and importing errors, as well as allowing me to be productive in my daily work.

Experience collaborating with others. I have collaborated frequently with over 10 principal investigators from regional universities, state and governmental agencies, which has helped me to understand the diverse needs of different organizations and to find solutions which we couldn't resolve on our own. I have a manuscript accepted with revisions to the Journal of Integrated Pest Management, with ten authors. Working on this publication has required communicating and collaborating with all of the authors to prepare, review and submit the document in a reasonable time frame.

I believe that I meet many of the preferred qualifications stated in the job posting as well:

Experience in instructional delivery. I presented topics on Support Vector Machines and Principal Component Analysis for a book club I attended about Data Science, AI, Machine Learning, and Deep Learning in R. I have tutored Spanish conversation classes. I taught a class at local 4-H club, organized activities and encouraged handling of live insects during the 6th and 7th Tallahassee Science Festivals, led activities at two NFREC-Quincy Arts & Garden Festivals as well as NFREC Agriculture Adventures and Ecology Field days

Experience manipulating regular expressions. I have used regular expressions and stringr in R to manipulate strings and match strings as part of routine data cleaning done while preparing multiple data sets for statistical analysis and graphing for each of my research projects.

Experience in data validation. I have experience validating data types in R, which has a number of built-in packages to test the structure of data frames. I have used these tools in each of my research projects to handle missing values and to verify that my data is in the proper format for statistical analyses.

Familiarity with agricultural cropping systems. My MSc in Entomology was from the University of Idaho, in an agricultural cropping system in the Pacific Northwest: I did research on potato psyllids and Zebra Chip with Dr. Erik Wenninger at the Kimberly Research and Education Center near Twin Falls, Idaho. My work there helped me to become familiar with the pests that affect potato crops in the region: I processed hundreds of yellow sticky cards, which collected a variety of insects of concern for Idaho farmers. I also lived for around 15 years in Moscow, Idaho, so I am familiar with many of concerns for agriculture in the region, especially northern and southern Idaho.

Previous experience working with databases. I have not worked with databases specifically, but I have basic understanding that R has some packages which integrate with SQL, and I have heard that the *_join functions from dplyr mimic some of the functionality of SQL merges. I am certain that I could quickly learn to integrate the use of a database into my workflow as needed.

Ability to work and solve problems independently. I have had multiple opportunities to work effectively with minimal supervision: I frequently had to commute between field sites for a field experiment with collaborators in Georgia. As a visiting student researcher, I would be gone for days working on independent projects as I conducted multiple surveys of roses throughout southern Georgia and northern Florida, staying overnight in local hotels as the need arose. I also collected data and set up my experiments for field trials in Georgia independently, and I would stay at student housing for multiple days while working on my projects. I also have experience solving difficult problems on my own, as I created a DIY method for field extraction of plant volatiles with inexpensive equipment from a hardware store. This was done in order to continue our research without moving infected plants across state lines, to comply with permitting restrictions. I also had to adapt our research when we became restricted from using roses for experiments on FSU campus, due to COVID-19: we found an alternative site for our samples and were able to conduct an entire field study at a nearby plot.

Effective verbal communication skills. I believe that I have the requisite verbal communication skills, as I have presented orally to various audiences, including scientists, stakeholders and the public at over 11 conferences, and club meetings. My experience with Extension has also allowed me the chance to communicate with the public in various ways, which has required me to adapt my communication style to fit the audience. I also speak, read and write Spanish fluently, which allows me to reach and communicate with Latin American audiences.

Remote candidates who are unable to work on campus must have access to a Zoom-capable internet connection. I would be glad to move back to Idaho, but I also have computers which are capable of remote work, and a stable internet connection.

In conclusion, I believe that my agricultural research and data processing experience make me a well-qualified candidate for this position. I have three enjoyable years of experience working with the staff of the University of Idaho during my MSc, and I would love to return and apply what I have learned from my PhD to wrangle agricultural data for UI.

Thank you for your time and consideration,

–Austin N. Fife