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 the past six years of my MSc and PhD in Entomology. I often clean and manage multiple files simultaneously: for example, I had recently had to clean and combine over 50 .csv files of 'dirty data' into a single data sheet for analysis via PCA and UMAP.

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 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 as 2 articles in peer-reviewed journals, a technical report, I also have a manuscript in review for the Journal of Integrated Pest Management. I have written over six abstracts for conferences, and I have written 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 the majority 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 is from the University of Idaho, I did research on potato psyllids and Zebra Chip with Dr. Erik Wenninger at the KREC near Twin Falls, Idaho. I became familiar with the pests that affect potato crops in the region, as I collaborated with the KREC staff and the Aberdeen lab. I also lived for around 15 years in Moscow, Idaho, so I am familiar with the challenges that agriculture faces in the region, especially in Idaho and eastern Washington.

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 remotely with minimal supervision: I was a visiting student researcher at the lab of a collaborator in Georgia, and I would be working alone for days while I set up and managed our field trials. I also conducted surveys of roses along the Florida-Georgia border, often driving alone for hours and staying overnight at hotels as needed.

Effective verbal communication skills. My six years of experience with Research and Extension Centers has allowed me to communicate with fellow scientists, stakeholders, and the public at over 11 conferences as well as more informal settings like 4-H meetings. I also speak, read and write Spanish fluently, which permits a wider audience to access our research.

Remote candidates who are unable to work on campus must have access to a Zoom-capable internet connection. I would be glad for an excuse to move back to Idaho, but I also have a stable, Zoom-capable internet connection here in Florida.

In conclusion, I believe that my experience with data curation, along with my experience with agricultural research in the Pacific Northwest make me a well-qualified candidate for this position. I have had 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