Basic questions

Agricultural Research Service scientists gather and maintain large amounts of valuable data in the course of their research. In order for the agency to provide better access to data storage in the future, please provide your answers to the questions as soon as possible. Focus on your most important data associated with research projects, the data you have used to generate publications, and especially the data you will need for future uses.

* 1. Your email address

[   ]

* 2. Research unit (mode code)

[▲▼]

[   ]

3. Are you answering this survey about your own data, not a group's data?

◯ Yes

◯ No (please specify how many people your answers represent, including you)

[   ]

4. What data types do you use or generate? (Check all that apply)

☐ Tabular (e.g. XLS, CSV) ☐ Image (e.g. JPG, PNG, SVG, TIFF) ☐ Statistical (e.g. SAS, SPSS, R)

☐ Database (e.g. SQL, MS Access) ☐ Video (e.g. AVI, MOV, MP4) ☐ Scripts or code

☐ Genomic (e.g. FASTA) ☐ Models (e.g. Matlab or other)

☐ Geographic Information Systems (e.g. KML, GeoTIFF, SHP) ☐ NetCDF or HDF

☐ Other (please specify)

[   ]

How much, where, and how often

5. Estimate the amount of research data you currently maintain. DO NOT INCLUDE back up copies or online public databases you manage.

	< 500 GB	500 GB to 1 TB	1 to 5 TB	5 to 10 TB	10 to 100 TB	> 100 TB
Total	○	○	○	○	○	○
Active data only (will be used next 6 months)	○	○	○	○	○	○
Inactive data only (stored for long term)	○	○	○	○	○	○

Comments

6. Roughly how many files does this represent?

○ 1 to 1,000 ○ 1,001 to 10,000 ○ 10,001 to 100,000 ○ More than 100,000

○ Other (please specify)

7. Where are these data currently stored? (Choose all that apply, again, don't include backup data)

☐ Hard drive of the instrument which generates the data

☐ Local hard drive (in your computer)

☐ External hard drive

☐ Shared network drive (e.g. research unit or location server)

☐ ARS data center or ARS collaborative platform (give name in "Other")

☐ Other (please specify)

☐ Ceres (High performance computing cluster)

☐ Cloud (e.g. Amazon Web Services)

☐ Data repository (give name in "Other")

☐ I don't know

8. How much research data do you anticipate generating in the next two years?

( ) ≤ 500 GB   ( ) 1 to 5 TB   ( ) 10 to 100 TB

( ) 500 GB to 1 TB   ( ) 5 to 10 TB   ( ) > 100 TB

( ) Other

[   ]

9. How often do you typically change (update or expand) your active data?

( ) Daily   ( ) Weekly   ( ) Monthly   ( ) Annually   ( ) It is relatively static

( ) Other (please specify)

[   ]

USDA Agricultural Research Service

SCINet and Ceres

SCINet is the network backbone to rapidly move files to and from Ceres, a High Performance Computer (HPC) that includes 69 nodes with 1560 cores and 15 terabytes of memory.

10. Are you or do you plan to be a SCINet user?

( ) Yes   ( ) No

USDA Agricultural Research Service

11. Have you loaded your data onto Ceres?

( ) Yes   ( ) No

12. If you have loaded data onto Ceres, how did you do it?

◯ Direct SCINet connection ◯ Over VPN ◯ Shipped hard drive

Other, or why not (please specify)

```


```


Final questions

13. Roughly what proportion of your data is already associated with published research (e.g. peer-reviewed articles or ARS-published reports)?

◯ 0 ◯ 10% ◯ 25% ◯ 50% ◯ 75% ◯ 90% ◯ 100%

Comment

```


```

14. How long is your data typically worth keeping? (Select as applicable)

	During the life of the project	Up to 5 years after the project ends	Up to 10 years after the project ends	More than 10 years after the project ends
Tabular data or database	◯	◯	◯	◯
Image or video	◯	◯	◯	◯
Scripts or code	◯	◯	◯	◯
Model runs	◯	◯	◯	◯
Geospatial data	◯	◯	◯	◯
Genomic data	◯	◯	◯	◯
Other (note in comments)	◯	◯	◯	◯

Comments:

```


```

15. Describe any special circumstances or needs for your data (e.g. THREDDS server, protection of personally identifiable data, etc.)

16. List URLs of any online public databases you host. (These should not be included in your responses above. Additional survey may follow.)

17. Do you have any additional comments or concerns? Please share them with us below.