




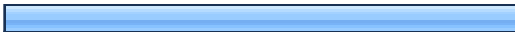














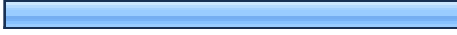











1. To begin, please tell us your primary role at Clemson.

		Response Percent	Response Count
Faculty		40.3%	129
Staff		3.1%	10
Graduate Student		55.0%	176
Other		1.6%	5
answered question			320
skipped question			1

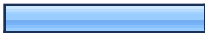





2. Which of the following best describe the types of data you have produced, or anticipate producing, as part of your research? If you are unsure, please select "Other" and provide a brief description of your data.

		Response Percent	Response Count
Non-digital text (hand-written notes/sketches/paper laboratory notebooks)		51.9%	137
Digital text		76.9%	203
Images		53.4%	141
Audio recordings		22.0%	58
Video recordings		30.7%	81
Spreadsheets		61.0%	161
Databases		37.9%	100
Computer Code		31.4%	83
Artistic products		5.7%	15
Digital gene sequences, or similar digital renditions of biological/organic/inorganic samples or specimens		10.2%	27
I don't produce any data		3.4%	9
	Other (please specify)		12
		answered question	264
		skipped question	57










3. Please select, and/or list, the file extensions you have produced, or anticipate producing, with your research.

		Response Percent	Response Count
pdf		91.2%	237
txt		52.7%	137
doc		88.5%	230
xls		68.1%	177
jpeg		64.2%	167
tiff		35.0%	91
m		8.8%	23
gif		20.0%	52
csv		31.2%	81
ppt		70.0%	182
mov		20.0%	52
wav		18.5%	48
html		25.8%	67
wmv		13.5%	35
aiff		1.2%	3
	Other (please list)		44
answered question			260
skipped question			61

4. What is the largest amount of digital data you have produced for a single project, or anticipate producing, in the near future?

		Response Percent	Response Count
1 GB (gigabyte) or less		29.8%	79
More than 1 GB but less than 100 GB		34.0%	90
More than 100 GB but less than 1 TB (terabyte)		12.5%	33
More than 1 TB but less than 100 TB		4.5%	12
More than 100 TB but less than 1 PB (petabyte)		0.4%	1
More than 1 PB		0.0%	0
I don't know		18.9%	50
answered question			265
skipped question			56




5. What strategies have you adopted to protect your data from corruption or loss? Select all that apply.

		Response Percent	Response Count
Copies of data sets are saved on a disk, USB drive, or computer hard drive		87.5%	230
Copies of data sets are saved on a local server		24.7%	65
Copies of data sets are saved on a central campus server		19.8%	52
Copies of data sets are saved on a web-based or cloud server		40.3%	106
Copies of data sets are stored in a data repository or archives		9.5%	25
Backup files are automatically generated		21.3%	56
Backup files are manually generated		29.3%	77
Access to files is restricted		22.8%	60
I have not adopted any strategies to protect my data from corruption or loss		3.0%	8
	Other (please specify)		12
		answered question	263
		skipped question	58






6. Please indicate how often each of the following statements is applicable to your data after completion of your research projects.

	Always	Sometimes	Never	Response Count
An appropriate version of the data (i.e. cleaned, de-identified, etc.) is made publicly available on my website, a publisher's website, or at a public data archive or repository	9.4% (22)	42.7% (100)	47.9% (112)	234
The data are not made publicly available, but I respond to individual requests from scholars on a case-by-case basis	22.1% (51)	56.7% (131)	21.2% (49)	231
The data are not made publicly available beyond the members of the research team	16.8% (38)	54.0% (122)	29.2% (66)	226
			Additional comments	15
			answered question	239
			skipped question	82

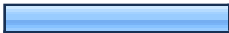


7. Clemson University Libraries will soon be rolling out an Institutional Repository, where Clemson's data can be collected, preserved, and shared in digital form. Would you consider submitting your data to this repository? Note that many funding agencies request plans for long-term preservation of access to data.

		Response Percent	Response Count
Yes		57.6%	141
No		7.8%	19
I don't know		34.7%	85
	Additional comments		22
		answered question	245
		skipped question	76




8. What might prevent you from sharing the data you have produced or intend to produce for your research?

		Response Percent	Response Count
Little value to others		34.4%	72
Confidentiality or privacy issues		64.1%	134
Commercialization or patent issues		34.9%	73
The data I work with has license or usage restrictions that prevent sharing		16.7%	35
Data requires secure access I am unable to provide		12.0%	25
	Other (please specify)		18
		answered question	209
		skipped question	112

9. "Metadata" refers to descriptive information or documentation about data. Have you produced, or do you anticipate producing, metadata for your projects?

		Response Percent	Response Count
Yes		33.5%	82
No		27.3%	67
I'm not sure		39.2%	96
Additional comments			6
answered question			245
skipped question			76

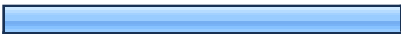






10. If you answered "Yes" to the previous question, does the metadata you produce conform to known standards in your discipline?

		Response Percent	Response Count
Yes		31.5%	39
No		6.5%	8
I'm not sure		62.1%	77
Please specify the standard(s) you are using.			9
answered question			124
skipped question			197





11. How important do you think it is for Clemson to spend resources on providing the following services?

	Very important	Somewhat important	Not important	Response Count
Short-term data storage/preservation (5 years or less)	58.4% (135)	37.7% (87)	3.9% (9)	231
Long-term data storage/preservation (more than 5 years)	43.9% (101)	49.6% (114)	6.5% (15)	230
Data security support	71.9% (166)	24.7% (57)	3.5% (8)	231
Guidance on depositing data into repositories or archives	60.0% (138)	36.1% (83)	3.9% (9)	230
Guidance on how to use appropriate metadata standards	43.6% (99)	48.0% (109)	8.4% (19)	227
Guidance on writing a data management plan	50.9% (116)	40.4% (92)	8.8% (20)	228
Guidance on intellectual property issues with my data	61.3% (141)	33.9% (78)	4.8% (11)	230
Guidance on privacy/confidentiality issues with my data	64.6% (148)	32.3% (74)	3.1% (7)	229
			Additional comments	9
			answered question	231
			skipped question	90

12. For which data management activities would you be likely to seek assistance (such as workshops, or one-on-one consultations), if offered?

		Response Percent	Response Count
Creating a data management plan		59.7%	129
Data security, storage, & back-up while research is ongoing		71.3%	154
Secure retention & disposal of selected data after research completion		50.9%	110
Sharing & publishing of research data for validation and re-use		52.3%	113
Archiving data & providing long-term access		56.0%	121
Producing metadata for your projects		31.9%	69
Compliance with policies, legal requirements, and ethical standards		54.2%	117
	Additional comments		6
		answered question	216
		skipped question	105

13. In your opinion, where should funding come from to cover the costs of data management and storage for research supported by grants, contracts, or other external sources of funding?

		Response Percent	Response Count
It should be folded into the direct costs of those grants and contracts by the individual researcher or research team		17.7%	40
It should be paid for by the University from overhead/F&A (Facilities and Administrative) funds it receives from grants and contracts		46.9%	106
It should be paid for by the University from other sources of funds		8.8%	20
No opinion		26.5%	60
	Other (please specify)		17
answered question			226
skipped question			95

14. Please share any additional thoughts or concerns you have regarding campus support for complying with funders' data management policies.

	Response Count
	14
answered question	14
skipped question	307

Page 2, Q1. Which of the following best describe the types of data you have produced, or anticipate producing, as part of your research? If you are unsure, please select "Other" and provide a brief description of your data.

1	Drawings, models and presentation boards.	Nov 16, 2012 11:48 AM
2	I'm not sure where survey data fits in this list but I collect a lot of survey data.	Nov 15, 2012 6:26 PM
3	Motion Data, 3D-Animations	Nov 14, 2012 9:45 PM
4	integrated mechanical and electronic devices	Nov 14, 2012 5:53 PM
5	poster plotter	Nov 14, 2012 4:09 PM
6	poster plotter	Nov 14, 2012 4:06 PM
7	journal papers, book chapters	Nov 14, 2012 3:30 PM
8	Survey data	Nov 14, 2012 3:27 PM
9	graphs	Nov 14, 2012 3:25 PM
10	Books, articles.	Nov 14, 2012 3:23 PM
11	GPS Data	Nov 14, 2012 3:23 PM
12	electromagnetic signal amplitudes in time over large areas	Nov 14, 2012 3:23 PM

Page 2, Q2. Please select, and/or list, the file extensions you have produced, or anticipate producing, with your research.

1	.ai .dwg .skp	Nov 16, 2012 11:48 AM
2	.sav .spo	Nov 16, 2012 11:00 AM
3	.vi	Nov 15, 2012 6:51 PM
4	wma (Windows media audio file) sav, sps, spv (SPSS data, syntax, and output files) mxd (ArcGIS project file)	Nov 15, 2012 3:56 PM
5	png mat mp4	Nov 15, 2012 12:55 PM
6	mp3	Nov 15, 2012 12:32 PM
7	.sav	Nov 15, 2012 10:49 AM
8	mp4	Nov 15, 2012 10:19 AM
9	none	Nov 15, 2012 9:09 AM
10	.dwg	Nov 14, 2012 11:31 PM
11	the motion data can for example come as v-files (Vicon), mb-files (Maya Binary)	Nov 14, 2012 9:45 PM
12	wpd; cdx	Nov 14, 2012 9:27 PM
13	mp4	Nov 14, 2012 9:19 PM
14	cpp c py ps eps	Nov 14, 2012 6:55 PM
15	I produce tex (LaTeX) files for papers, though the papers will often be published from the tex format to the pdf format.	Nov 14, 2012 6:38 PM
16	Adobe Illustrator Adobe Photoshop	Nov 14, 2012 5:40 PM
17	mp3 mp4	Nov 14, 2012 5:39 PM
18	edp (FreeFEM)	Nov 14, 2012 5:37 PM
19	.psd .indd	Nov 14, 2012 5:30 PM
20	mp3,mp4	Nov 14, 2012 5:20 PM
21	png	Nov 14, 2012 5:18 PM
22	psd ai indd	Nov 14, 2012 5:14 PM
23	par	Nov 14, 2012 4:38 PM
24	.sav (SPSS)	Nov 14, 2012 4:17 PM
25	mat	Nov 14, 2012 4:12 PM
26	.png	Nov 14, 2012 4:10 PM

Page 2, Q2. Please select, and/or list, the file extensions you have produced, or anticipate producing, with your research.

27	cel bam sam vcf fasta fastq dat genome	Nov 14, 2012 4:08 PM
28	dta , sav, por (Stata and SPSS)	Nov 14, 2012 4:06 PM
29	R, plt, dia, fig, tex, bib	Nov 14, 2012 4:04 PM
30	SAS	Nov 14, 2012 4:03 PM
31	vi (LabVIEW)	Nov 14, 2012 3:52 PM
32	tex	Nov 14, 2012 3:44 PM
33	png, cpp, hpp	Nov 14, 2012 3:42 PM
34	opd	Nov 14, 2012 3:36 PM
35	netcdf	Nov 14, 2012 3:33 PM
36	.tex (TeX files)	Nov 14, 2012 3:33 PM
37	.cdx .png	Nov 14, 2012 3:32 PM
38	fasta, gff3, sam, bam	Nov 14, 2012 3:31 PM
39	Acc Mp4 Pages Keynote Numbers	Nov 14, 2012 3:31 PM
40	.dwg, .psd, .ai, .xml	Nov 14, 2012 3:23 PM
41	.mat	Nov 14, 2012 3:23 PM
42	dta is a big file extension we use in the Business School	Nov 14, 2012 3:21 PM
43	mat, dae, sdf	Nov 14, 2012 3:20 PM
44	xml	Oct 5, 2012 7:36 AM

Page 2, Q4. What strategies have you adopted to protect your data from corruption or loss? Select all that apply.

1	Backed up automatically one one external hard drive while i manually back up to a different external drive that is kept in a fire safe off site.	Nov 16, 2012 11:48 AM
2	I also use dropbox for non-sensitive items.	Nov 16, 2012 10:03 AM
3	Hard copies	Nov 15, 2012 11:25 AM
4	Have my own RAID server	Nov 15, 2012 10:19 AM
5	multiple external hard drives locked in safe	Nov 15, 2012 6:30 AM
6	Data are generated on one computer (e.g. the computer running an instrument) and analyzed on a different computer, thus there are always two copies (of the raw data, at least).	Nov 14, 2012 8:57 PM
7	A university solution for backup has been impossible to arrange. CCIT wants several thousand dollars per year for off site backup. Grants can't stretch to this. Yet this is a critical function which I believe F&A should cover.	Nov 14, 2012 7:46 PM
8	Time machine backup, and self-maintained RAID backup drive	Nov 14, 2012 5:40 PM
9	I could use help in this regard. I tend to back up my data and projects when I remember. I work from home and on campus. I often email documents as a way to back them up. Some data I have stored on the Palmetto cluster and I rely on the university to do backups. I would like help in setting up some sort of systematic backup plan. Some data I must restrict access to for privacy reasons, other data are not restricted, I just don't want anyone to mess it up.	Nov 14, 2012 4:03 PM
10	dropbox has become very useful	Nov 14, 2012 3:58 PM
11	external drive	Nov 14, 2012 3:44 PM
12	Paper copies of data are saved.	Nov 14, 2012 3:43 PM

Page 3, Q1. Please indicate how often each of the following statements is applicable to your data after completion of your research projects.

1	This was a very confusing question - I'm not sure what you mean by a "research project" - like there are multiple major projects, but not mentioning studies?	Nov 14, 2012 10:18 PM
2	Ideally, I would make the data publicly available. It's a question of having the time to do it.	Nov 14, 2012 9:48 PM
3	As I am funded by NIH and NSF all data must be published as open access.	Nov 14, 2012 8:01 PM
4	For "Always" read "whenever possible". For "Never" read "Not unless it's unavoidable."	Nov 14, 2012 7:02 PM
5	I have not yet created usable data, hence my use of Never above, but I plan to make the data I create available to individual bases and not to the general public.	Nov 14, 2012 6:41 PM
6	I have not yet had any requests for data from scholars, but if that were to happen in the future, then my answer above would be applicable.	Nov 14, 2012 5:46 PM
7	Unless published, access to data is strictly controlled to research team members	Nov 14, 2012 5:42 PM
8	Not very sure of these past 3 questions/answers yet.	Nov 14, 2012 5:17 PM
9	poster session	Nov 14, 2012 4:08 PM
10	These categories are confusing. Sometimes we release cleaned, deidentified data to a public repository. I haven't yet, but others in my group have, I think. Some raw data remain identifiable and are never made public. We have some collaborations where I might respond on an individual basis, and pending consent, release select data on a case by case basis. My projects are not federally funded and funding is less than \$500,000.	Nov 14, 2012 4:04 PM
11	Raw data gets converted to publications-so some of it is public	Nov 14, 2012 3:37 PM
12	The answer for these questions depends on the project nature and the sponsors terms	Nov 14, 2012 3:30 PM
13	Data is trying to be made public, but caught in red tape.	Nov 14, 2012 3:24 PM
14	I haven't engaged with these funding agencies	Nov 14, 2012 3:24 PM
15	I wanted to post my data at my department's website years ago for scholars to access. I asked the chair for permission and he refused to allow me to post the data.	Nov 14, 2012 3:22 PM

Page 3, Q2. Clemson University Libraries will soon be rolling out an Institutional Repository, where Clemson's data can be collected, preserved, and shared in digital form. Would you consider submitting your data to this repository? Note that many funding agencies request plans for long-term preservation o...

1	Thank you. I've been hoping this would happen.	Nov 17, 2012 8:11 PM
2	depends on how the data would be shared	Nov 15, 2012 9:23 PM
3	It depends on the cost and the degree of security and control I have over access by others.	Nov 15, 2012 6:29 PM
4	There are IRB considerations.	Nov 15, 2012 4:01 PM
5	I am not presently funded, so I pretty much do my own thing. If I had an NSF grant I would do it.	Nov 15, 2012 11:15 AM
6	I guess that would be up to my PI	Nov 15, 2012 10:32 AM
7	I'd need more information before making this decision.	Nov 15, 2012 6:32 AM
8	Sharing research data is extremely harmful to IP protection	Nov 14, 2012 10:18 PM
9	Seriously? Do you have any idea about the amount you are talking about and the cost? It could be more than 1 PB on day one and grow from there. You are over interpreting the directive for long term storage of data by funding agencies. Central databases for genetic data, for example, already exist. It would be mad for CU to try to perform these functions. You need to be very clear about what data you are going to store. Storing everything isn't an option.	Nov 14, 2012 8:01 PM
10	I think this depends on the access available to non-Clemson users.	Nov 14, 2012 6:41 PM
11	I will not upload any data unless it is published.	Nov 14, 2012 5:42 PM
12	libraries should be working with ccit's data storage architect on this topic.	Nov 14, 2012 4:57 PM
13	Depends on supervisor's opinion	Nov 14, 2012 4:44 PM
14	My concern would be data sharing	Nov 14, 2012 4:43 PM
15	Only if restricted access options	Nov 14, 2012 4:25 PM
16	This would actually enhance my grant applications	Nov 14, 2012 3:37 PM
17	This would be very valuable!	Nov 14, 2012 3:32 PM
18	Depends on how confident I feel about my data	Nov 14, 2012 3:24 PM
19	Assuming that I'd have the flexibility in accessing the data, and amending the data	Nov 14, 2012 3:24 PM
20	I would give my data if they wanted it.	Nov 14, 2012 3:23 PM
21	Sounds like a cost, with no benefit. Offering backup facilities in exchange sounds interesting.	Nov 14, 2012 3:21 PM
22	This would be very helpful!	Oct 5, 2012 7:38 AM

Page 3, Q3. What might prevent you from sharing the data you have produced or intend to produce for your research?

1	Much of the data I use is governed by publication agreements and it is only available to people who specifically receive permission. Also, data developed using federal dollars can be archived at ICPSR which is the best known repository of social science data. I would probably store it there if I wanted to make it generally available to other scholars.	Nov 15, 2012 6:29 PM
2	Copyright issues would prevent some from be shared for other than classroom use.	Nov 15, 2012 12:35 PM
3	HIPPA issues	Nov 15, 2012 12:00 PM
4	nothing	Nov 14, 2012 9:50 PM
5	only the time needed for cleaning everything	Nov 14, 2012 9:48 PM
6	Until the work is published it is proprietary. Once it is published raw genetic sequences or databases of images are available for the data to be published to. This is more valuable because the data will then be in context with like studies. Standalone storage by CU will not be very useful unless the search engine to find info is practically google. And even then the effort to upload and categorize the data so that it is searchable is unattainable.	Nov 14, 2012 8:01 PM
7	I try to avoid any of these issues.	Nov 14, 2012 7:02 PM
8	I am not doing any research.	Nov 14, 2012 6:27 PM
9	It typically is for a very targeted group.	Nov 14, 2012 5:44 PM
10	not sure of the libraries ability to handle load requirements.	Nov 14, 2012 4:57 PM
11	I might not be done analyzing it all yet. I am mostly concerned about the privacy issues.	Nov 14, 2012 4:04 PM
12	the time/effort needed to describe data sets does not excite me	Nov 14, 2012 4:00 PM
13	I don't mind sharing my data but I'd like to know who has it. Hence the case-by-case approach. Also indiscriminate data sharing may adversely affect my ability to be "first in press". Since I'm paid by Clemson to publish, that's not good for my career.	Nov 14, 2012 3:57 PM
14	Doesn't apply. I am a scientist and always share my data.	Nov 14, 2012 3:38 PM
15	I think there are important concerns about sharing data prior to publication.	Nov 14, 2012 3:32 PM
16	Nothing would prevent me from sharing data	Nov 14, 2012 3:24 PM
17	The work involved.	Nov 14, 2012 3:21 PM
18	Some of the data is too far away from utility. We try to thin it down to what is needed to reproduce studies.	Oct 5, 2012 7:38 AM

Page 3, Q4. "Metadata" refers to descriptive information or documentation about data. Have you produced, or do you anticipate producing, metadata for your projects?

1	Not 100% certain what metadata I need to have, but it is needed.	Nov 15, 2012 11:34 AM
2	Does that include data resulting from analyzing sequence data? if so, then yes.	Nov 15, 2012 10:32 AM
3	Too time consuming, and needs a lot of manpower.	Nov 14, 2012 8:01 PM
4	I would do so if I were aware of standards for providing such metadata.	Nov 14, 2012 7:02 PM
5	I have not in the past, but I believe a project I plan to start on in 2013 will move in that direction.	Nov 14, 2012 5:44 PM
6	Several of my colleagues have had discussions re this-metadata seems to depend on who is defining it. A publication could be metadata. So its hard to answer	Nov 14, 2012 3:37 PM

Page 3, Q5. If you answered "Yes" to the previous question, does the metadata you produce conform to known standards in your discipline?

1	Well, some of it will. Some of it will be crap. Heuristics can be fun like that.	Nov 15, 2012 10:32 AM
2	Standards do not exist	Nov 15, 2012 10:20 AM
3	We use what seems standard, but we aren't following a particular protocol for the metadata--just what makes sense to us.	Nov 14, 2012 8:59 PM
4	Probably not, since I am unaware of any standards	Nov 14, 2012 5:22 PM
5	Possibly not any standards in my discipline.	Nov 14, 2012 4:58 PM
6	its structured for my use only	Nov 14, 2012 4:00 PM
7	ASTM and ISO	Nov 14, 2012 3:38 PM
8	embedded in netcdf	Nov 14, 2012 3:38 PM
9	personal tagging / group folksonomies	Nov 14, 2012 3:36 PM

Page 4, Q1. How important do you think it is for Clemson to spend resources on providing the following services?

1	We need to be a 21st century research university.	Nov 17, 2012 8:13 PM
2	I have not gathered data yet so it is difficult to say.	Nov 15, 2012 2:00 PM
3	I believe it is up to individual researchers to provide storage for data and not necessarily the responsibility of Clemson.	Nov 15, 2012 6:35 AM
4	I don't require guidance on IP or privacy/confidentiality because I restrict all data to lab members until published. Short-term and long-term storage would be helpful but only if access was restricted to members of the research group with ABSOLUTE certainty that no other person outside or inside the University could access the data.	Nov 14, 2012 5:51 PM
5	please be working with CCIT.	Nov 14, 2012 4:59 PM
6	Most of these services are important, because they are and will be part of the requirements for funding proposal.	Nov 14, 2012 4:19 PM
7	The questions are difficult to answer as they are worded. Specifically I think short & long term data storage are very important, but I can't and won't depend on Clemson for that. I do it myself.	Nov 14, 2012 4:01 PM
8	I believe the long-term data storage is a high priority because this is difficult to do with the resources from externally funded projects, that are inherently time-limited.	Nov 14, 2012 3:45 PM
9	We are far behind in this area, but doing even half the things on the list could put us far ahead	Nov 14, 2012 3:39 PM

Page 4, Q2. For which data management activities would you be likely to seek assistance (such as workshops, or one-on-one consultations), if offered?

1	My professors usually help with all of the applicable elements in our data management.	Nov 15, 2012 2:00 PM
2	I do security research and teach security. Would be interested in working with you on those issues.	Nov 15, 2012 10:22 AM
3	None of the above.	Nov 14, 2012 8:18 PM
4	Creating a data management plan is a must.	Nov 14, 2012 5:25 PM
5	I heard about penalties related to HIPAA violations, now I think the risks to inadvertently releasing any kind of identifying data far outweighs whatever benefit to sharing data. I can't keep any of my research data on my laptop because of that. A couple years ago I took a CCIT class on backing up your data, but it was more along the lines of bring a DVD or email a document to yourself or pay for some sort of service. I wish we had a system where I could easily back up my laptop. I wish I knew how to encrypt it or something like that. I would be willing to pay for a backup service. I haven't done it because I would have to research it and wonder if they are legitimate or what the right procedure would be. If CCIT had a link I could click and I could just have my laptop automatically backed up just like I get regular software update notices, I would like that a lot.	Nov 14, 2012 4:18 PM
6	Don't need help as of now.	Oct 5, 2012 7:42 AM

Page 4, Q3. In your opinion, where should funding come from to cover the costs of data management and storage for research supported by grants, contracts, or other external sources of funding?

1	Perhaps a combination of option 2 and 3	Nov 15, 2012 6:54 PM
2	SOME BALANCE BETWEEN THE FIRST 2 OPTIONS WOULD BE BEST	Nov 15, 2012 12:02 PM
3	Under no circumstances should direct funds be used for this! This is another example of administration gone mad. We don't need another CU policy that takes time away from research or money away from the actual science.	Nov 14, 2012 8:18 PM
4	Some combination of the above if necessary, but primarily from overhead. That funding should cover research generated without external funding as well.	Nov 14, 2012 7:04 PM
5	Because each researcher has different needs, the researcher should make the choice. Direct costs are the logical way since a researcher that does not require the above described services would not have to budget or pay for them. Those that did require the above services, would budget appropriately. However, forcing all researchers to take a portion of the direct costs to fund this service is illegal.	Nov 14, 2012 5:51 PM
6	I do not believe there is one option here. I have students working on special projects that are unfunded and would need the support. Then there are grants that I think certainly should contribute to the expense.	Nov 14, 2012 5:47 PM
7	Or other funds	Nov 14, 2012 5:25 PM
8	The University should pay for it from other sources, just like the library and CCIT. If you charge for it, people will figure out work arounds that are free and non-standard and bad. If I had had to pay for my time on the Palmetto cluster, I never would have used it or developed any expertise in it. I think if you try to fund something like this that takes real expertise and needs to be done right and will cost some real money, the expense needs to be spread out, not taken from the few people who manage to get big grants. You could charge the grantees a fee for specific data submissions, but the infrastructure and training and maintenance costs need to be borne by the University as a whole.	Nov 14, 2012 4:18 PM
9	Could be either F&A or other sources. There is not room in already tightening direct costs.	Nov 14, 2012 4:12 PM
10	Direct costs would be a very bad idea. It's often difficult to access what amount and type of data will be produced. So, who to do the costing for that? Furthermore, the university provides hardware support. As I see it, data should be at the same level and not treated differently.	Nov 14, 2012 3:45 PM
11	Most granting agencies WONT cover the cost of data storage-that's one of the things indirect costs is supposed to cover!	Nov 14, 2012 3:39 PM
12	(I hope the order of this list was randomized for each user)	Nov 14, 2012 3:27 PM
13	But I'm not sure with NSF grants, and what they fund.	Nov 14, 2012 3:26 PM
14	I would emphasize that it should come from Overhead- Administrative funds it already receives from grants.	Nov 14, 2012 3:25 PM
15	If the University benefits from the data, then the University should fund it. If the	Nov 14, 2012 3:23 PM

Page 4, Q3. In your opinion, where should funding come from to cover the costs of data management and storage for research supported by grants, contracts, or other external sources of funding?

University does not value the data, then the University should have individuals who value the service pay for it, be it from grants or out of their own pocket.

- | | | |
|----|---|---------------------|
| 16 | For projects that generate exceptionally large amounts of data, the costs should be included in direct costs. For "standard amounts of data (e.g. <5Gb/year) costs should be covered by overhead. | Nov 8, 2012 9:02 AM |
| 17 | The University should be committed to storing knowledge in a mineable way as should government repositories and journals. | Oct 5, 2012 7:42 AM |

Page 4, Q4. Please share any additional thoughts or concerns you have regarding campus support for complying with funders' data management policies.

1	Thanks for focusing on this important issue.	Nov 17, 2012 8:13 PM
2	Needs to be on a very secure server	Nov 15, 2012 12:02 PM
3	It is unwise to allow access of data to entire university. Not sure what the difference is between this proposal and what is currently offered by many departments (network storage accessed by username & password specificity)	Nov 14, 2012 10:22 PM
4	There seems to be a confusion in the questions here. Are you trying to provide backup for research data, or are you trying to put the data online for the whole world to see. We produce stunning amounts of data and it would be nice to have a backup for that so that we don't accidentally lose the information - we already take local measures to protect the data. However, we don't want this data to be available for the whole world. The standard in my field is that raw data is kept for five years post publication. Large data sets are deposited to centralized databases following publication so that other researchers have access. I think you are over interpreting the data sharing policies of funding organizations. The publication of the data and submission of data to centralized databases is all that is required. I review large numbers of grant applications and have never seen the interpretation you are attempting to derive from the data management policy of the funding agencies. CU simply cannot manage or afford the amount of data storage that you appear to be proposing.	Nov 14, 2012 8:18 PM
5	I am not sure of the motives with this plan. Since I take care of my own data management, I strongly do NOT approve of funds from my direct or indirect being taken to fund this type of service. Only those that require the service should pay for it. Because current CCIT costs are so high, I doubt you will convince many to participate.	Nov 14, 2012 5:51 PM
6	The funding should be from whatever funds the libraries. Libraries store things like dissertations and whatnot and this is fairly similar in my opinion.	Nov 14, 2012 5:26 PM
7	Would CCIT (Campus Tech Support) be involved or would this be a separate management group? How many physical locations of access/management would there be?	Nov 14, 2012 5:24 PM
8	Some of us will be doing research that is part of our job but not supported by grants, contracts, or other external funding sources, so we would need access as well, and Clemson should pay for that somehow.	Nov 14, 2012 5:00 PM
9	My comments may have been general and not necessarily about compliance with funders' data management policies. I don't currently have any big dollar federal grants supporting my research. I'm unclear on whether you were asking about data management as a repository for things like memos, emails, training manuals, procedures, documents v. raw data that could be used for analysis by others.	Nov 14, 2012 4:18 PM
10	I think it is important, but the acceptance will depend on who is paying for it. If the \$\$\$ from direct costs, I would rather continue to copy my data on an external disk that I then put away.	Nov 14, 2012 3:45 PM
11	If the university is going to claim indirect costs from granting agencies, then	Nov 14, 2012 3:39 PM

Page 4, Q4. Please share any additional thoughts or concerns you have regarding campus support for complying with funders' data management policies.

providing data management capabilities is a moral imperative.

12	Spend university resources elsewhere	Nov 14, 2012 3:36 PM
13	Great initiative, hope it is implemented soon!	Nov 14, 2012 3:28 PM
14	If we are going to a a player in ultra data-rich technology, we're going to need PB of free storage. In genomics, I am already limited by the world-class HPC storage resources at Clemson. I still haven't decided what I'm going to do long term.	Oct 5, 2012 7:42 AM