my name is Tom Melvin I work here in the

Morris library I'm a librarian

in the reference and instructional

apartment I liaison to most not all but

most of the engineering departments on

campus as well as marine studies geology

and geography so I have met some of you

through that

thanks for coming I want to give you a

little background about what we're doing

today about needed i/o let me back up I

will be presenting with the Nita Swartz
from IT client support services and we

bottner from the research office the

three of us have been working on this in

various ways for a few years now about

we can't decide four or five years ago

I'm gonna say four to make that yeah but

I'm younger if I say four Anita and I

took part in a workshop sponsored by the

Association of college research

libraries which was a month-long online

workshop dealing with data management

plans and that was our first
introduction to this and then a person
here in the library Greg silvus and I

took part in the East Science Institute

about three years ago which was a

slightly more comprehensive overview of

research data management which

culminated in a two-day conference out

in Dallas so we had gathered a lot of

this knowledge and about a year ago it

looked bout a year and a half ago we

decided to offer a workshop to make use

of the knowledge but also we want to see
what the interest was on campus this had been a topic that had been coming up a lot we did not know how much interest on campus there would be so we decided to offer a workshop on an introduction to research data management that was offered last October a year ago last October and we had a really good turnout and we offered it again in March again had a good turnout and we had always intended to branch off and do some more specialized workshops after that so
Anita and I got together with Lee and decided to go ahead and do a specific workshop on writing a DNP we're basically extracting what was a small part of our original workshop and stretching it out into a large a workshop I don't think it was anybody here at any of the other workshops we did I looked at the program list I know that's good because we do repeat a little bit not a lot but it all be new to you that'll be wonderful
so if you have any questions as we're going along do not hesitate to ask this will not take the complete out well take about an hour will not take the whole 90 minutes we're leaving plenty of time at the end for questions or if you need help registering to use the DMP tool anything like that so we'll have plenty of time at the end so what will we cover we're going to take a look at a research guide that I have created on data management research guides are built on
a platform called Lib guides that we use here in the library it's basically the easy way to create web pages very easy to update I just want to make sure you're aware of what's on there we'll talk about locating information on funding agency requirements so that you will know where to go to find out what the foundation or funding agency that you are dealing with requires in terms of a data management plan or if they need one at all we'll talk about
locating sample data management plans so that you can find some that have already been done take a look at them see what you have to do we'll look at some directories for both metadata schemas and data repositories both of which are things that most data management plans will want you to include so we want you to know where to go to look for these things we're going to have a fairly extensive look at the DMP tool which is an online tool that not only will walk
you through the process of creating a
data management plan but it also has a
lot of great resources that you can use
to prepare so we're gonna look at both
the resources and the process of walking
through to create a plan we'll talk
about data management budgeting
including funds in your budget request
that to take care of your data
management it does cost money like
everything else and we'll talk a little
bit about data data attribution and some
of the ethical legal issues involved

again if you have any questions do not hesitate to put your hand up or way to the end either is fine the first thing

we want to point out is the research office that we are presenting with does have a wonderful page if you go there you will find a lot of good resources

and in fact I'm going to back out of here for a second and actually go there the one we really want to point out is the researchers tool
so here is the home page and if you click on where it says for researchers you will then go into what they call the researchers toolbox and there are a lot of great links over here on the right on conducting research funding opportunities a grants management guide the intellectual property guide which will take a look at it a little longer so we just want you to be aware that there's a lot of great things here they do link out to my data management card
from here as well so be aware it's there

it's a good resource we want you to know

that there are places to get started so

keep that in mind

also as I mentioned I have created a research guide I do not expect you to recall this URL we just put it up there

so I'm going to show you how to get there research guides are part of our platform for the library and I'm not sure how many people are actually aware of them so if you are at the library's
homepage if you use this much you will have noticed we have what I like to call an icon based search system so whichever icon is highlighted is what you will be searching when you use the search box the default is always our catalog delicat discovery so whenever you have originally go to the page that's what you're going to see but you could search to see if we have a particular database or in particular a journal if you want that's what those will do I'm going to
go to the research guides and if you want to see if there's a research guide for any particular subject that you're looking for just type in whatever subject that is business economics philosophy and it's just throwing a keyword search through the research guide so if there's one there will come up I'm just going to type data and when I search I will get a number of links to my guide which I will then bring up so this is a guide I have built it is
constantly being worked on I would welcome any suggestions if you ever have anything that you think would be useful to put up there please email me I'd be happy to put it up it's a work in progress and in many ways by nature it's general it's supposed to as a place to point you out to places to get started so the first page is just a basic introduction to the concept of research day and management there will be slides from our previous
workshops and we will be putting the slides from this workshop up there so all the slides you see today you will be able to get to I'll have them up by tomorrow or maybe even this later this afternoon some good sites for general information on research data management some basic reasons of why you want to manage and store your data you can contact me of course then we'll have sections specifically on the DMP there will be a link out to the
DMP tool which we'll be taking a look at later. A basic checklist of things that you would like to include on a data management plan. The basic structure there will be links out to sample data management plans. We'll take a look at some of these later on and some basic templates for what is required depending on agencies and things like that. Then a page on metadata what it is why you need it some very good metadata directories that we'll
take a look at some of the more common metadata schemes that are used and some general information.

same thing with file formats we will not be talking a lot about this today need a we'll touch on it we've kind of taken some of the very technical stuff out of the previous presentation to concentrate more on the DNP.

but there is information here choosing the right file format open versus closed data and then there'll be a page for
data repositories with subject specific repository guides basic insulin safeguarding your data the concept of open data and a link to UD space which is our institutional repository here at the University which we'll talk a little bit about so again we just want you to be aware that this page is there again it's a work in progress I'm always changing it it's very easy to update it's one of the wonderful things about Lib guides
so again if you run across something that you think would be useful go ahead and send me a link and I will include that in here.

Okay now this is a chart we included this in the previous workshop and but it's handy and the reason we like to have this is to really emphasize the point that thinking about your daily management should occur throughout the complete research lifecycle it's very typical for us to think about it at the
end we've done our research we have this data I need to keep it somewhere but you really need to start thinking about it early on and this emphasizes that so when you're planning when you're identifying the grants of funding will the person the entity you're dealing with require a DMP if so what elements do they want in it describing you know organizing your assets it helps you get organized who will be taking care of the data throughout the research process is
someone responsible for backing up the

data where will it be kept so you want
to plan these things in advance
certainly when you're doing the research
it helps you organize your assets
describe the assets and analyze the
assets again if you've planned in in the
beginning for this that helps when you
publish you will identify open access
publications hopefully it'll help you
share and cite the work where you're
going to deposit the work are there
limitations or restrictions discovery

and impact when people start using your data and finding it in various places

there are databases now in fact we have access to one data source which do nothing but index datasets just like an index for journals you can now have an index for datasets so your data set can be cited just like a Journal article that helps with your citation counts so it will certainly be used then and then the traditional idea of the preservation
of the date where are you going to keep it long term do they have any restrictive policies how long are you going to keep it up things like that so the whole point here is that you need to be thinking about this throughout the research lifecycle I want to stop for a second and go back William do you want to talk a little bit I had just talked to talk a little bit I had just talked about ud space this is my colleague William Simpson
and he is the what is there an official title here the coordinator of UD space

okay okay so he's gonna take a few minutes just a tumblr I'm doing space

okay so yeah Judy has an institutional repository yes it's a server computer

it's housed in this library okay so it's on a secure network I have some pursuers that'll give you know policy information

a little bit more information about on the side counter over there bumps in my business card my contact information is
on the back depending on which department or Research Center you're in you may already be participating and not realize it so first thing I would do is just go to the institution repository locally its branded as UD space never gifts how we came up with that name so you can find it that way the documents that are in it are searchable by Google history repository does have its own search engine but if you search we also find them that way either Google or
Google Scholar last year the Faculty

Senate approved for the first time an open access collection so for our peer-reviewed journal articles there is a designated collection in the institution repository you go there move you know if your community or a research center is not participating still have a place to put your research articles and you can do that yourself is actually an online forum where you submit the information about the journal article
and then the library staff takes it you don't have you know and a person responsible for uploading our girls and it is optional and it's not all your journals just free the information with a page because it tells you what publication B basically we just started last year and then going for so we don't want things from in 1920 anybody have a quick question about the institution poster I don't want to take a lot of time so mainly I just want to
plant the seed in your mind that we'd
have one had one place for a long time
that's professional advantage any
questions about if you're interested in
participating you know if you're here as
a department chair I don't know everyone
here but if you're here is an apartment
here or so it takes their information
back and recommend it to your director
or department chair I'd love that it
would be great basically we get the word
out like this and workshops a word about
speaker that we'll talk about so I was looking at it and it seemed like it's mostly used to our pilot reports articles and things like that certainly the organizers Club that's all purity ring's now is it also can it be used for say just archiving it's required you know that's that's a great question and that's one that's comfortable the answer is not the best one but it's is it it really depends it really
depends on how much data you have and

then also whether once the date is put in the repository whether you want to actively use it okay so if you're just

archive it and if they don't want private format it's and if you're repurposing it to go back you know if you're gonna be using it every day or every weekend that might not be the ideal spot for it and if it's not we can

Tom our IRA you can recommend another yeah any weeks so it really depends
100 so ya get a grand had it resubmitted

with this and identified the repository

yet and I'm still waiting you're back on

my shirt wait lady worked it out and his

data management plan was to say that you

would like a representative sample was

dated in the institution but instead of

Walden he wasn't going to be he's going

to Boston

okay alright that's a good example to

use anybody else have a quick question

okay later said the brochures again over
there on the side counter you can take

that and you can contact me if you have

any questions okay so this slide we include this because a lot of people say

well why is this so important now it's kind of come up in the academic world a

lot the whole idea of data management so

this is a brief timeline I'm not gonna spend a lot of time on this but I think

it's interesting to look at

so about 13 years ago when over 13 years ago exactly now the NIH started
requiring that any application grant

application for over half a million

dollars required a data management plan

or you had to state why it was

impossible do one I have no idea how

many people actually just stated why it

was impossible and I imagine quite a few

and it stayed that way for many years

and then in early 2011 the net the NSF

required all grant applications to

submit a two-page D at P and when they

say two-page DMP they mean no more than
two pages one of the things we really want to emphasize today is that this does not have to be a large onerous process they want a little bit of information they want you to think about it write it down but most of these data management plans you look at or one or two pages long so keep that in mind that was followed a few months later by what was called the research works Act which was an act put into Congress backed pretty much by publishing concerns
who are worried about losing control

over their copyright etc so it was kind

of a push back and it really went

nowhere but it was then I answered by

the federal research Public Access Act

which also went nowhere but which was

the first act presented which really

addressed the issue of making publicly

available federally funded research to

the public so neither of these really

went anywhere they both died in

community in May of 2012 there was a
petition put up at whitehouse.gov the

require free access over the Internet to

scientific journal articles arising from

taxpayer funded research petition and it

generated the required was at twelve

thousand or twenty five thousand I'm not

sure hits signatures whatever that then

responded the White House responded and

they responded in February 2013 with an

OSTP men room from the Office of Science

and Technology Policy which I'm holding

here in my hand
entitled increasing access to the

results of federally funded scientific research and they set out the policy

principles and what's interesting to us in our context today is that this was the first document that talked about data specifically before this it has been access to the articles itself in here there is an objectives for public access to scientific data in digital format so this is the first time the federal government actually addressed
the data there was then an executive order put out by the White House thirteen six four two entitled making open and machine readable the new default for government information and it as well contained a section I believe it is section three implementation of the open data policy so that set the wheels in motion federal agencies had to start getting their ducks Lenora in a row about how
they were going to deal with this and

the first one to come out was Health and Human Services in February of 2015 where they had their processes and requirements lined up and other agencies have been coming online since then so that's why you're hearing so much about this now it has become pretty much deira Gore that federal agencies will require you not all do yet we'll talk about how to find that but you should always be thinking when you apply for a grant am I
going to have to include a data management plan with this okay so why is this important well as we talked about a little bit earlier it helps you stay organized throughout the research process you will comply with requirements either of the journal or the funder which is what we're really concentrating on today it will certainly help with your ability to validate reproduce your research as well as the ability to disseminate your data for
other people's use and you will conform

with legal and ethical responsibilities

and I'm gonna spend a minute talking

about this last one not a lot of time

but we need to touch on that so what are

the legal and ethical issues

well the biggest one is who owns the

data and it's very important that you

realize what the policies are of

whatever institution you're at regarding

who owns the data and I'm going to Bop

back to the research office page right
now and if I go into the intellectual

property guide right here you will see a

link to various types of IP and there is

one here for data and I'm not going to

read the whole thing but I will do the

highlights at the end here data obtained

by faculty staff or students in the

course of employment at the university

or work directly related yada yada yada

is property of the University data

obtained from the use of tangible

resource materials in research at the
University legally belongs to the University now that's a starting point

I'm sure there's many and and I guess

Lee could tell talk about this a little

bit later that there's it's it's not hot

and carved in stone but that's the

starting point so realize that when you

were doing research here at the

University and this is true of most

institutions they're going to claim

ownership of the data it's important to

know that yes
I'm bored you push this to leave okay

yeah yeah good question interesting okay

so that's one of the things you want to

keep in mind

oops excuse me how can the data be used

so one of the things we'll talk about in

a minute is attributing a license to the

data when you put the data up you might

have restrictions on how you want it

used you might want attribution for its

use you might not care at all there's a

couple of different organizations that
have created some standard licenses that you can mention either in the data management plan itself or on the site where you store the data or both so be aware of that the issue of collected dated versus the system used for storage is simply the idea that your data the collected data might have different ownership and access issues then the system you are using to store it on so when you choose a repository to put your data in you should be aware of what
their policies are if there's any restrictions and that goes for does a database have copyright protection that's normally not the case but if they have massaged the data in any way they might claim unique use to the data so that's again something you need to be aware of so these are issues that usually don't come up a lot but you certainly want to be aware of them especially the ownership of the data I mentioned the licenses so the two most
common groups the first is the open data Commons group and they've come up with three basic licenses and you can just attribute your data one of these licenses the first is the public domain dedication and license which basically says it's free to use for everyone to use as they see fit you know they don't want attribution you don't want any kind of credit just take the data and run with it the second is an attribution license
where users are free to use the data

bass and its content in new different ways provided they provide attribution
to the source of the data and/or the database so at that point you're just saying we'd like attribution the third is the open license database which basically is like the attribution database but an unrestricted version of the new product must always be accessible and any new products using that data must be distributed using the
same terms so it's a little bit more restrictive so again if you want to when you're in talking about data repositories in your DMP or when you actually put the data up you can say we would like you to comply with the open data Commons public domain or attribution or open database the other group is the Creative Commons they have some licenses the CCO licenses basically like the first one no attribution required and the public domain mark was
used for me up Marc works that are in

the public domain and for which there

are no known copyright or database

restrictions so you can flag factual

data in a database in order to make it

clear that it's free to use so this goes

back to the issue of or if you're in a

database that might have any

restrictions you can say yeah but my

date is fine don't worry about there

there are issues any questions about any

of this okay I'm gonna let it need to
take over

oops nope not yet I'm sorry I forgot I stumbled so we're gonna talk a little bit about the first steps you want in managing your data the first thing you're gonna want to think about is the criteria for keeping what the data so you're getting a research project you're generating lots of data you might not need to keep it all you first thing you have to do decide is what do I how much do I want to keep what is going to be
important for keeping that's a decision

you're going to have to make based on

your research needs you will want to

consider the metadata that you're going
to be using to store the data we'll talk

about metadata in just a second and then

you're going to want to think about

where you want to store your data both

in the short term during the research

process and long term for preservation

so what I'm going to do is pop over to

the data management plan
and take a look at some of the resources which will help you with this now.

there's a link to the DMP tool right off my page or you can just google it up and Anita will be talking about how to log in and things later on for what I'm going to show you are some resources they have that will help you plan ahead and you don't even have to log into the site to do that.

so when you get to this home page you can go to the resources and training.
section and just click on view all and then you'll see two links here at the top DMV requirements and public DMPs that's what we're going to take a look at the DMP work whoa just mad let me back up for a second I got ahead of myself I'm sorry we're gonna look at metadata first because back here in the slide if you recall number one you will take care of number two we'll talk about metadata so off my page if you go to the metadata
section it'll talk about what metadata is basically metadata it can be defined as data that defines data it's the data that will be used to locate your data and the nice thing is that different schemas have been created for certain subject matters so it's not something you have to invent and what it basically will do is say these are the elements that we need to be included in your metadata when you put your data up and that often will vary from subject to
subject so on here you will have a few

minute data directories that I want to

show you the Society of American

archivists has a good site it's a but

the page doesn't display well I don't

know what is going on here and it's not

a browser issue I've tried this on

various browsers various computers but

they just have an alphabetical list here

of metadata schemas by certain subjects

or categories so categories for

descriptions of works of art and you can
just go through here and see if there is

one that matches what your needs are

I find the digital curation Center to be

a little

more well-organized again incident

alphabetical order

agricultural metadata schema astronomy

visualization metadata so again you can

just browse through here see if there's

a metadata schema that will fit your

needs

and as a librarian I don't normally push
Wikipedia but they've actually got it very nice it's not the enemy it's very it's a good starting point but they've actually got a really good page here and they do a very good job of describing what metadata is and then they've got this nice subject listed list based upon what area you're dealing with so for instance you probably don't even know this but in librarianship we use the machine readable cataloging metadata system so all this is all the records
you see and delq at our catalog using

the mark records which basically attributes a field number to every element of the bibliographic record

title author publisher and year

published whatever so it gives us a logical way of cataloging that information so when you go into the system you can find it all so that's the metadata schema we use in the library world I'm pointing it out just because being a librarian so just be aware
they're out there depending on the complexity of your research and the data you're using that will have a lot to do with the metadata schema you pick you can do that you can just come up with your own you do not have to pick another one I will also list here on the page some of the more common metadata schemas used so again just be aware that it's here and you can use it okay I think now those of the manage of directories I
just showed you and now sorry Anita will
talk about various places long term and
short term where you can keep your data
welcome everyone so far does anybody
have any questions before I start moving
into the next part great so where should
you keep your data Tom heads or we've
had part of the presentation talked
about the throughout the research
process and so we are trying to
basically give you some guidance to talk
about what options you might have that
is during the actual project itself and then storing it long term so I know this sounds as though it might be as Tom said not the best answer in all of the cases but we because of us being as part of the university we have IT professionals within different colleges and within different departments you would probably be best to start there to ask your particular IT professional exactly what departmental storage or backup solutions exist they
may not be exactly the same from college to the college and department department or perhaps you're collaborating with an institution so you might want to talk to those again your IT professionals and see who you're collaborating with to see what options already exist and that IT central IT I'm part of IT client and support services but we work with network and system services and we provide central services on campus IT services we may not have a central IT
service that exists specifically for you

for your data storage or backup however

we're always available to consult with

on specific research needs especially if

it's something that's in the part of the

designing the grant or setting it up

that this can be budgeted for if you

have special circumstances that need to

be met and those special circumstances

are based on the type of project or the

type of research that you're doing and

understanding the requirements of your
data things such as the size

specification or is it sensitive data

that has to be encrypted or other types

of things that have to be done with the

data in order for it to be accessible by

whatever methods

supposed to be long-term shared I have

the speed of the drives that's really

based on the performance how often do

you need to access it is there high

speed issues involved is it a takes a

long time to upload something and store
it but you're not really going to need to access it all the time so are those things acceptable for your project and then last but not least Tom talked about this a lot regarding the licensing or the grant restrictions is it something we've been talking mostly about open access about the data that's supposed to be shared you need to understand the implications about your specific grant and whether or not you can use one of these open access repositories or
locations for storing them and that's

also true of UD space is that not
correct William UD space is based on it

being open access is that correct

okay so again if you have restrictions

that are associated with that particular
data or part of what you're doing for

your research then you may not be able
to use those particular repositories so

we're talking remember about storage

throughout the process and then beyond

the long-term storage so central IT does
have some storage with very specific limitations or restrictions so you'll want to read about what that means for your actual project the win domain file services are based on being able to set up file shares or services that could be used within your department and usually we can ordinate those with again your IT professional they set up and arrange to be able to get some type of disk space associated it's usually a mounted type of file system that you
would have available on say your Windows

or your Mac system your desktop or

perhaps in your lab machines we also

have a an agreement with Google Apps for

Google Drive that there's an unlimited

amount of storage but again it depends

it's a cloud-based system the speed at

which you can upload and download files

may not be

suitable for your needs and as a result

it may not comply with your ability to

put those files up in the cloud and so
careful consideration should be taking no matter what cloud resource you use before you're storing the data on a non-UD server and so this link and we are going to make these slides available because I've provided a lot of links here goes to our IT security page and talks about those considerations what you need to think about with respect to where you put the data up in the cloud as far as backup services itback backup services again this is a link to the
site that talks about it in much more
detail but we back up our data and email

that are we are part of all of our

central systems and we provide recovery

so that means that your email and

perhaps other central services that

you're using like the wind DFS and some

of those other file system services we

make sure that you can recover files or

that they're backed up or that they're

part of our disaster recovery so that's

part of our central services and then we
also have other fee based services that

you can set up for servers within your

department and again usually that's

coordinated with an IT professional you

would talk to them about providing the

perhaps the servers that you're using

for collecting data or storing data for

your research in a lab you might want to

make sure that that is set up to be

recovered as part of this recovery plan

and that's based on how much data and

how much storage is required that
you want to backup and that you pay for

that basically but you can schedule it

and it's not that expensive and it's

worth it to be able to do but it's for

department own servers only so again

it's not something that we're doing for

personal desktop systems I did add this

CrashPlan personal for students and

employees again check within your

department because some of the colleges

and departments have set up a department

or
why'd CrashPlan or some other type of mechanism for backing up your personal laptops your desktops we've talked about you know the worst case scenario with respect to your research is that your laptop is stolen or one of your devices that you're using to collect the data is destroyed you want to make sure that you can always get a backup of that so if your department or college does not provide that directly for your desktop systems or maybe some of these smaller
systems that you have in your lab then

you can consider using this CrashPlan personal which is for students and employees here and it's a discounted rate I think it's around 25% on the plans that you can get that you can set up to do it yourself

any questions about throughout the lifecycle of the project as now we're just going to switch to talking a little bit about the long term data storage yes

okay so now are you by any chance in
Health Sciences yes school nursing okay

so specifically in um I don't think is

anybody here from the biostatistics oh

you're from IRP okay so you're from IRP

so I was just I was specifically asking

about Health Sciences because they

actually have several staff that are

really helping people with their data

specifically in Health Sciences Ryan

polling is one of the persons and

William I can't think of his name he and

I can give you that information before
you leave today they're helping to

basically figure out the types of things

about how to set up and plan for the

data what you have to change in the in

the long term about that what you need

to plan for even when you're doing the

the statistics or whatever observations

or any type of work that you're doing

with that how you should plan that that

should be stored in order so that there

isn't any identifiable information in

there you have to basically have a
separate set of the original data and

then the data that you’re using for the

actual research so they can definitely

help you to plan about that they give

some examples on their website about

that as well and I thought they might be

here today to introduce them but I can

definitely pass their the information on

to you so that you can get some yeah

exactly and he has like on his website

he has a little thing that talks about

the file formats and what you have to do
with the basic information about how the
the keys and you know between a
worksheet that describes each of the
values and then how that translates into
the data that to give it meaning after
you do the analysis so pre analysis and
post analysis so yeah they have that up
there but they and they definitely said
that they're helping people do this
because it's part of their grant funding
as well they get part of their salaries
and things like that are based on
helping with individual grants good

question any other questions

okay so the long-term storage we've been talking about that with respect to the

preserving and sharing why is it important you might have read articles from research that hasn't been done 25 years ago and they're unable to reproduce the results from that research

they can't find the original datasets

they can't even find the original software perhaps that was used to
generate that data so they can't even go back and reproduce that which is much of why a lot of this the mandates and much of the legislation has come out because of the fact that they want to try to allow people to build on pre-existing research do you don't have to start all over again you want to be able to use where somebody left off and take that further so basically it's all it all stems from how you're going to be able to share that preserve it and share it
so that somebody else can use it, it can be found. It can be then taken for yourself to reproduce and then move further to get those new discoveries from the original data so in most of the data plans it asks you will your data be accessible and usable when it is needed and they want you to be able to answer those questions so a data repository may or may not include the preservation, and the preservation there may be things such as the file formats and how you
change those file formats when you

preserve the data maybe you’re using

Excel for most of the data but you might

want to just store the final results in

a text file for example a CSV file comma

value separated for your preservation so

that you can then tell somebody all you

need to do is to import that into an

Excel spreadsheet or whatever else

you're going to use and you can

reproduce the results but the file

itself is an attached to a specific
software application or something that may limit what it's considered as far as being preserved the data repositories the funders that you have whether it's for a journal or a particular grant may in your in your subject may have a repository that's specific to you and you'll see that there are a lot of domain repositories we'll look at that shortly we talked I guess already about you D space and William talked much about that
it is as the question came up primarily

it seems for journals as far as articles and reports and things like that but they're looking more to expand on that so again on a case by case basis you know you should contact the library contact William if you have something that can be meet the needs of UT space and use the institutional repository we want to build on those relationships and try to see what kind of services we can expand upon in UT space so choosing the
right depository we already talked about

who can access the data and is it open

data and most of the repositories that we're looking at are based on the fact that it's open data are open access we understand at least what we've been told

is only about 10% of the grants or the types of research that people are doing

or restricted in the sense that there may be some time frame before it can be made accessible there may be a patent

that's going to follow or some type of
invention and that it can't be released

it may be working with a particular

corporation and they have restrictions

on that as to when it can be made

available so you have to just understand

that but in the case of the open data

then you can use one of these open

repositories this open data which let's

see oh and the data site so the reason

this data site and it says this re3 data

org they at one time were two

independent organizations data site it
kind of collaborated and put all of the services together under one umbrella called data sites and their goal is to be able to help you find and locate we use the data and also the aspect of the sighting like where where you can find the information about a sighted data source and then actually be able to use that data so the data sight basically now includes the R III data is a place where you can find the data repositories but the data site itself is
how you can find data and related data

yeah I was that I forgot how you said to do that so okay both of those sites will be on yeah I think that one the research guide under data repository yes and you will see them there right so data site here and then here’s the R III data org so in in the data site you can get to everything as it was before you can they’ve really changed the site quite a bit now it looks really nice it makes it much easier to get here but again under
resources things like the metadata

scheme all of those things that are linked to from the research live guide

much of the work that was put into that live guide we had originally collected

data and information from the classes

that we took Tom and I and then from other sites and other people that we

talked to from their research data

management plans and how they were basically getting information out to the

University so again the services here
through the site guide it tells you each
one of these steps but if you go to the
our III data here's where you'll find
all the data repositories and then it
links to the site itself and again each
of these sites are based on subject
based oh I meant to point out before I
go back very quickly the guide
guidelines on the open access and
scientific publications and research
this is again another place where you
can go to to refer to information about
open access and the guidelines referring to that but on the AR III data site again you can just search here and find I'm just gonna say engineering for example it's supposed to bring up basically all of the related repositories that are associated with engineering as an example are there any questions yes if you have a publication and you're required to then you know give people access to the raw data these
repositories mm-hmm are there ways to

sort of I like you know that this is

document yeah and actually this this list

that we have here on the repositories

this one is specifically by subject or

discipline and so if you have if the

journal actually most of the journals

now are saying if the data has to be

published with the paper that's that

whole idea is that within that data site

it's allowing you to be able to

basically site and link the data with
either that journal or as itself so that

if you're in your journal you refer to

the data you may be using one of these

sites and it is actually cited with that

paper it's part of that paper and that's

the whole purpose of these organizations

to make sure that you can find

everything no matter how it's associated

with the research good question and

thank you I forgot to bring up this link
to the repositories here by subject

which is a really good place as well
so I think that concludes my portion and

we're going to go into why we need to

have the actual data management plan

take it away Tom so now I jumped ahead

earlier but we're gonna take a look at

some of the resources to help you get

started with agency and foundation

requirements and some sample DMPs

so just as a real quick review why have

this you want to meet the agency

requirements very important hopefully

this will increase your citation count
as data sets themselves become things that can be measured

certainly ensure the research integrity

and replication facilitate the sharing

of data help manage the data and allows you to budget for data retention which we'll talk about so for number one I had jumped ahead with this but now is the right time if we go to the DMP tool and just go again you don't have to log in to get to these basic resources go to the resources and training section and
then you will see a link to DNP

requirements and public DMPs these are

the two were going to look at right real

quick so when you click on the DMP

requirements and i am just going to

bring up all of these you will get a

list of templates the template will say

what is required in the DM polly by the

funder be it a foundation or an agency

it will have links to the funder and

then there will be some sample plans I

have to be honest the sample plans that
you see here are not very good so I do not advise that you really use them.

I will show you some better places so you can just go through here and see if the Department or the funding agency that you're looking for has something up and what I'm going to do since my background is engineering for a librarianship we're going to take a look at the NSF engineering template which I have a copy of right here and I also have right here so here is the basic template and one
things you'll notice is the template itself is a lot longer than the actual plan has to be there go into pretty good detail of talking about what they want so they will list the roles and abilities of the people involved in the project as far as data management goes they will want to know what the expected data is what data do you expect to collect what data do you expect to keep they will want to know the period of data retention how long will you be
keeping it formats and metadata that

you'll be incorporating your policies

for dissemination and public access will

there be publication delays as a Neto

mentioned a very good reason might be

you're applying for a patent you don't

want to publish this information yet so

you'll put an embargo on some of the

information for awhile and then the data

storage and preservation of access so

what you should do is go through this

list and see if you can find your
foundation or agency if you do not you
can certainly just try to google it up
you do not have to go to this list if
you don't find anything I would strongly
urge you to contact the funding agency
that you're dealing with and see if they
do have any requirements and if so what
they are don't assume that they don't
want the data management plan just
because you can't find it out you don't
want to have to go back and start over
again
then they have these very nice public DM

P's up here and these are data

management plans that people have

actually put together using the DMP tool

and have actually submitted and they

have said sure go ahead and put my plan

up so again they'll have the plan title

the template that it was based on the

institution the person's from and the

owner so what I did was I just went

through here and let's view them all

again and I went through to find one
based on the template that I was using

and found a researcher and here is that

data management plan itself it is for a

research project called inquiring into

engineering faculty assumptions about

students a profound human systems

intervention and what they base that

we're going to do was interview

engineering faculty about their students

so the whole date management plan is

right there again does not have to be

long so what I'm going to do and I've
shortened this down to make a little bit easier I'm going to go through and I have summarized the template and then taking an example from that data management plan so we're just going to run through these six requirements real quick to see what they did in this case give you an idea basically how easy it can be so for the roles and responsibilities who will be delegated who will monitor the
data management plan and who will have responsibilities over time for the decisions about the data so from their example the primary responsibility for managing the data will belong to the PIL is the Schlemmer if the p.i is not available both Linda van supa and Aaron Estrada will be able to access the data and continue to management now they've got more here but that would have been enough the rest they're talking about where they're
going to store it which comes later they

put it in here redundant so again two

sentences and they've met the

requirement next section expected data

describe the types of data the samples

the physical collections software

curriculum materials describe the

expected types of data to be retained as

opposed to all the data you'll be

collecting and describe the file types

data set size number of affected files

file sets content so the data for this
project will include 60 interviews with engineering faculty and a set of survey data the interviews will be stored in both audio files and transcribes text files the surveys will be stored in Excel spreadsheets and include survey answers and demographic information transcribed text will be verified to the audio files no fewer than two copies of the data would be kept at all times and will be located in different locations so they're talking about backup security
in addition data for public access will be anonymized and a shift of the identifying information this process of anonymizing will be documented okay and that's about the most lengthy one in here period for data retention describe the period in a retention any exceptions of hiring longer retention periods and estimate how long your data will be kept after the completion of the research and their very brief data will be available for ten year period after
completion of the grant which seems to be a rather common period I'm wondering why that is but I run into that a lot data oh yes that's a good question and I have to say I don't know do you have any information on that I mean I believe it's published with each repository so yeah I mean as far as what how long it's supposed to be there and how long it's you know that is actually included as part of that why they wanted in the DMP itself so you know if it's not going to
meet those requirements then you know it

you can't use it obviously but

in most of the cases that we've seen for

e.g. I forgot to mention that we are

members the University of Delaware's

members of the icpsr and so when you use

icpsr they specifically say how long it

will may be maintained in a certain form

of accessibility okay once it's in the

repository then it'll say it'll be

stored with whatever the guidelines that

icpsr has for beyond that particular
point and that's what people write into

the into their DMP and then if the grant

gets accepted based on that if the grant

gets denied and says you have to do

something else beyond that point then

that might be where you have a special

circumstance where it has to be moved

after a certain period of time and again

that is described in the data plan the

data management plan

for ten years and he comes with an

expiration there they actually taken it
down deleting it I mean what happens to

the audience is still that response only

give that the I who may not even be know

right yes that's good question I don't

know and I would bet it would be the

responsibility P I but that's just a

guess but I'll look into it that's a

good question yeah that has not come up

thank you okay next data dissemination

and policies for public access sharing

in publication delays articulate how

sharing a primary data is to be
implemented the policies for public access and provisions for protection of privacy how will day to be shared and managed with stakeholders and clearly state publication delay policies if there are any again so it is the goal to house the data in the digital Commons at Cal Poly that is their equivalent to UD space the data will be uploaded at the completion of the grant the individuals privacy really ensured by removing any identifiable information the data will
be publicly available through Cal Poly's digital Commons and through a research website the access to the data will comply with the IRB requirements although we will publish findings based on the data it is hoped that the data will be a rich source of inquiry for anyone interested in educational research we only request attribution in the use of the data so they're not claiming a particular license but they are asking for attribution so that's
where you would state that por say

create a commons license etc data

storage and preservation of access

describe the physical and cyber

resources and facilities used for

retention the long-term strategy for

storing archiving and preserving data

what archives repository database will

the day to be deposited in and what

procedures do they have for the

preservation of backup they being where

you are storing it so there example the
data will be housed long-term on Cal Poly's digital common with all the guarantees afforded data creation on our campus it is hoped that the data kept for 10 years after completion of the research and I would assume in this case if you William do you know what the policy is if for is for you D space for data retention okay data management and budgeting and so Lee's gonna take over on this data management is the budgeting for the data
management something you want to take into account when you are doing your grant it does take money we have had instances I know you might talk about this where somebody wanted to keep their data for 10 years somewhere and that was going to cost money but the grant would only pay for it for three years so what do you do then so these are things that need to be taken care of well goes to one place and it's open access anyone could get it right right right exactly
right so I just have a few quick slides

on budgeting I think as Tom has mentioned and even Anita as she’s
discussed they’ve been going through the slides

and talking about preparing a data management plan the idea of whether or not you can pay for the cost on a grant

is one that typically we get questions

from the audience so the answer is since

NSF was sort of one of the first federal agencies to sort of get on this

bandwagon of requesting data management
plans the answer is is that they have

guidelines for what is applicable and

how you actually budget and request the

funds so the according to their

frequently asked questions yes you may

request as long as it's in coordinates

with applicable cost principles it's

necessary to implement the data man

plan is an obvious and it must be

included online g2 on the proposal

budget and justified in the

justification that would be typical of
any expense that you have for NSF so

there's really nothing different cost

principles that they're referring to is

that the expense for data management

must be reasonable and necessary

it must be allowable under the

regulations or award provisions which it

is through NSF it must be allocable and

verifiable and that gets to that

question that that other audience member

had that Tom just mentioned on what do

you do if you need to budget you the the
solicitation is asking you to keep the data back around for ten years but the grants only three years so the answer is that it must be a proportional cost so we actually have to allocate a portion of that to another expense category on campus it the other thing with the cost principles is that everything must be consistently treated so I'm not going to read through because it's getting late and we're getting to the end of the time and we want to make
sure we have lots of questions but just like any other expense that you have you have to follow the cost principles and when it comes to data management expenses like I said at the beginning line g2 is specifically on the NSF guidelines where you should budget that line is actually the publication documentation and demonstration line and it's under other direct cost so if you have costs for storage and preservation or indexing that you need to actually
recoup through the grant you would

budget it on this line so then one of

the question becomes is there any

guidelines that other federal agencies

have put out and the answer is no but

does that mean that you can't request it

and the answer is also no you can

certainly request as long as data

management expenses are legitimate for

the the proposal and actually you can

walk through the cost principles as

appropriate for this particular
application then you can certainly put down the cost and request it from the federal agency okay yes it would go under g2 if it's an NSF are you doing an NIH because I heard you said yeah NIH other direct cost yeah questions about budgeting like I said I had three very short slides so I didn't have much to go on how many of you in the audience are actually faculty members here at UD have you written grants where you've requested data management cost no yes
did the funder fee pay for it and where

you funded you didn't put any yeah next

time do so you didn't either yeah we

actually it's interesting the reason I

asked the question we haven't seen a lot

of the budgets that are coming in to be

approved in the research office we have

not seen a lot of people requesting and

I think a lot of that is people don't

know that they can request so I just

wanted to raise your awareness that it

is something that you can do because it
is an expense that is now being required

of you and if they're asking for a data management plan so other questions how

can you - are there some guidelines for figuring out with this

just like you would do in any other expense you're gonna have to talk to IT

and you know whoever the end user or however you're going to have to purchase the services unfortunately we don't have any like set fee schedules or anything like that other questions
I think that's it so now we can actually move on has anybody here use the DMP tool at this point to write their plans yes and what did you think of it was it easy to use or it was easiest to use well we'd like to actually hear about that what was frustrating about it and because I mean obviously this was a tool that we basically in talking to other universities set decided to set this up because of the guidelines and the examples that they had so if as
we're going through if you can basically highlight and comment that as we're going through where which parts were frustrating got you okay yeah so that's probably not so it's not such a good situation but the idea behind the DMP tool is that you can actually when you sign up for the DMP tool it allows you to be able to bring in other people and collaborate on the DMP tool as well so that was one of the reasons why we thought that it was a good tool as tom
has mentioned about the resources periodically if I wanted to point out on the main page about the DMP tool news they're constantly updating this information to put you know basically like for example this one the Department of Energy has their generic template but then they have a new you know the NIH genomic data sharing they specifically put one in there for that now I took this snapshot a long time ago so when we go in and look live at the tool you'll
see that obviously things have changed

since that point as well it's very easy

to use by just clicking on either the

get started or login once you're there

and you can find this DMP tool by the

way it's on the research live guide it's

on the IT research computing webpages

it's you know we're linked to it from a

number of places I think even on the

research office office website as well

the key is is that we set up the

University of Delaware as an
institutional login which means that you can use your net ID and password to be able to get in. You do not need to create a separate account when you first log in. Once you select University of Delaware, it takes you to our authorization here at the University of Delaware for you to actually log in at this point with your net ID and password. It takes you to the part where you're going to set up your actual account.
in this case I'm just using a training account because I figured that this was generic enough but it shows you the basics using a you'd led you account you can put your first name last name the actual email and then it automatically comes up and says that tooth populated as part of the University of Delaware you also see this next section about an orkut ID how many people here have an ork ID great so there's some cards over there if you want to take this the
beauty of the orc ideas it sets up a
unique a persistent digital identifier
that distinguishes you from every other
research and it's not associated
with you by a specific institution it
becomes your unique research ID and that
way as you move or you change or you're
collaborating there's never any issue
about say who John Smith is at the
University of Delaware there may be 10
John Smith's but this auric ID then
allows you to be able to be that unique
and again places like the site the data

site and the the repositories having

doesn't help them to make sure that

eybody is being accountable and that

they're the persons who they say they

are so again in the DMP tool you can

register with that and here's a little

bit of information about the orc ID we

have this in the slides and again

there's some papers over there if you

like a little card about that if you

haven't had one I would highly
recommend setting it up so once you complete the profile again this is just you know options deselect or select in this case
I've just shown that I've selected everything which it's basically giving you prompts and options to tell you about the process as you're going through the DMP tool once you're in the DMP tool its referred to here as the the actual home screen or the login at this point and it basically has you know your
dashboard you can always go back under my profile where the location 4 is and make any of those changes if you want to go back and create your org ID or do all those things you don't have to do that at the very beginning and then of course you can get your DMPs or you can create the DIA the new DMP so I think Tom showed before about the DMP requirements once you're inside and you're logged in again you're gonna see these same templates that he referred to
before but now when you're in the DMP tool you may be using one of these to start out as the actual portion of what you want to use so if we want to go and select he was referring before to those examples meaning somebody already has done one and if you find one that is similar to yours you can select that and kind of use that as the base and modify it whereas if you’re starting with the actual template itself which is what I'm going to show you you're going to the
specific funder location and you're going to find out you're gonna start out with a blank plan in this case I'm showing under NSF part of the plan that Tom showed was I believe it's gonna when you look at the guidelines in the template it'll be obvious what I'm talking about but I'm just showing you two different ones a generic NSF and one that is specific for the social behavioral economic sciences so when we look at the generic NSF you'll see it
says types of data produce data and meta

standards and I believe that this

followed more of what Tom's sample

showed which just had the types of the

data the data and metadata the policies

policies reuse and then the archiving

and preservation whereas if you look at

the social and behavioral economic

science oh no I think this is this is it

because they had surveys okay they had

surveys and they had other types of data

this includes that roles and
responsibilities who's going to take care of the data so it actually did fall more under this particular guideline the expected data the periods of retention I think this was yeah I just I said that and then I said something else so yeah so this this was the pretty much the format that of the example that he had so that means if you could start out with that sample plan already filled out and then modify it if you think that a lot of that's the same and you could
change cally Paul to University of Delaware and things like that

the names obviously but if you go through if you go through the template

itself it's going to give you the same guidelines in the in this in this little section over here it basically tells you what it's expecting and then in the end when you go through the whole plan you can preview it and it basically shows you the plan itself builds it either in a Word document format or a PDF now the
idea is is that okay well you could just

do this yourself right let me go back so

you could just do this yourself by using

a word document and people say okay well

I could use Google Docs and then people

could share that way there probably

isn't any reason why you couldn't take

this plan and if that works better for

your sharing and collaborating then by

all means do this the nice part about

the plan is that it's a place for it to

be kept and reviewed and you can go back
and it has all those things like if you're done a section and you share this with somebody else at the universe that you're working on and you can also share with other people outside the university it's a little tricky if they're not members but as long as they have an ID and EMP tool then you can direct them to sections and it they can be like oh you need to do this section and you can coordinate that all within the DMP tool so the idea is is that it's
more of a sort of like I guess the way

that they do those grant online where

you're submitting the online grands and

they're asking you to go to sections and

review and add different text spires I

think is one of them in the government

where you're actually feeding in data

and different people are being queued

and prompted to do it so DMP tool is

similar in that sense that you can share

it but again there's nothing really

that's special about it other than you
can get these templates and resources

the templates and resources you can get

in Doc formats so that you could then

potentially use other collaborating tools does anybody have any questions

about the DMP tool this is just one way

to get people to not feel too different

that it's too difficult to create these

plans and now I guess in the end did you

end up doing a word doc yourself yes

so she prompted you go here and get the

template at least to find out these
sections for it okay yeah so I would say

that the googling might be a good way

but I would definitely make sure for all

of you that at least reviewing the

templates that are provided by DMP tool

they are reviewed by the institutions by

these granting agencies and we are are

told that these are being reviewed and

updated all the time so I would suggest

if anything that if you go here and as

Tom said you don't necessarily have to

log in to see the templates and the
tools and the resources but again you

don't really have to do anything if

you're using UD as the institutional

login you're setting it up one time and

it's kind of owned by you and then can

be tracked by you

as well any other questions I'll put up

our little thank you for coming today

and again tom is going to put the slides

up on the research guide you should be

for this afternoon

also you will receive an email with a
very brief survey if you could fill that

out we would appreciate it probably

later today or tomorrow if there’s other

things that we could add in in the that

would be helpful for these type of yeah

workshops if there’s anything that you

think that we should look at in research

to be able to provide other workshops

specifically or even if in your

department you would like us to reach

out and specifically address the

department's we’re more than happy to do
that from specifically the data

management plans but also you know reach out to IT if there's other things that you need with the IT professionals in your college and Department about budgeting we're really glad that we got that on the the workshop today we thank Lee for specifically coming because we always get questions about that and it's good to know that people can budget this any other questions thanks for coming thank you very much for coming