

Introduction to Research Data Management

Aims of course



What is research data?



"representations of observations, objects, or other entities used as evidence of phenomena for the purposes of research or scholarship".....includes:

- source data all data collected, created and used by the research, including data held elsewhere
- assembled datasets data extracted or derived from the above
- referenced data any subset of the above that has been used in analysis or to draw conclusions. Consistent with whatever is considered 'supplementary material' to research findings in your domain."

'Categories' of research data

What is research data management?



The "organisation of data, from its entry to the research cycle through to the dissemination and archiving of valuable results."



Making the case for Research Data Management





'Representation'

<u>Reversible</u>

Reproducibility of research











Genomic signatures to guide the use of chemotherapeutics



Excerpt from Keith Baggerly's lecture



Why is RDM important?





Thinking about data collection

Methodologies and standards



Taxonomies & controlled vocabularies

18-day pregnant females	", female (phenotyne)	bexanloid female	dikan/on	female_6_8 weeks old	
2 yn old ferro	ale. fermale.(pi	agnani) individual (iamale dioacious	siemala iemala, o	ther remate.
na an Arthur an Arthur	and the second	a and a start a	in an	<u>andra an an</u>	and the second
114	f	female, spayed	sex famala	408 yr old iamela	female child inn
stafamala	fancala	female, vingin	sexual-oxipanous iemale	aduli: ferinate	iemele mice 👘
ana ang ang ang ang ang ang ang ang ang	and Rassand and the second second	Constant of a maximized system of the	all in the track	constantial discontribution	$\label{eq:constraint} \begin{split} & \text{for each many many strength} = \frac{1}{2} - \frac{1}{$
sterile ismale		female plant	normal femories	<u>jismal</u> .	itemateriswonexteand
and the second s	Same and Same		Consulta - <u>Consulta</u>	<u></u>	<u> </u>
10.5. ST. 10.	Anna an Anna Anna Anna Anna Anna Anna Anna Anna Anna An Anna Anna	entre en la companya de la companya	a a suite ann an	ala, alashir	n se an
hitisalaana ka Sanay ka falasta	erenda erendaren er	the correct and a set of the set		Carriella di Ba	www.else.as.ldlever.udia.low.org.goglicy.
remale	female (calf)	femele	thelytoky	dioecious female	female, other
$(\mathbf{r}_{i})_{i} = (\mathbf{r}_{i})_{i} \frac{\mathbf{r}_{i}}{\mathbf{r}_{i}} \frac{\mathbf{r}_{i}}{\mathbf{r}_{i}} \frac{\mathbf{r}_{i}}{\mathbf{r}_{i}} = (\mathbf{r}_{i})_{i}$	$\frac{\partial \mathbf{r}}{\partial t} = \frac{1}{2} \left[\frac{\partial \mathbf{r}}{\partial t} \mathbf{v}_{1} \right]$	(a) <u>a series</u> (a) a series (a) <u>a series</u> (b) <u>a series</u> (b) <u>a series (b) a se</u>	an the state		$\label{eq:constraint} \left\{ \begin{array}{c} \mathbf{f}_{1} & \mathbf{f}_{2} & \mathbf{s}_{2} & \mathbf{f}_{1} & \mathbf{s}_{2} \\ \mathbf{s}_{2} & \mathbf{s}_{2} & \mathbf{s}_{2} & \mathbf{s}_{2} & \mathbf{s}_{2} \\ \mathbf{s}_{2} & \mathbf{s}_{2} & \mathbf{s}_{2} & \mathbf{s}_{2} \\ \end{array} \right.$
er versje oare de J	in and tarra	an da	A (1997) (1997)		anatar in

-

사람이 많은 것이 같다. 것이지 것이 많이 한다.

Nucleic Acids Research

이 사람이 나는 것을 알았는 것을 수 있는 것을 수 있는 것을 하는 것을 수 있다.

Over to you (DMP checklist)



Storing, describing, depositing, citing datasets





Naming your files



An example netCDF data file name is depicted below



An example from LSHTM

Funder requirements











RVC RDM polaVMCID BDH((abr)(Bdg)(ed))]TET /PA





Good alternative to



Backing up



Cloud storage





Publishing

Depositing

Studies may share their data by asving their data collection (or a subset) at a discipline based repository like the UK Data Acve (<u>www.dasa-arve.ac.uk</u>), or at an institutional repository that can preserve data and make them available to users.

Sequences must be submitted to sEMBL Database Library or GenBank. Protein sequences that have been determined by direcsquencing of the protein must be submitted to SWISS-PROT at sAll accession numbers should be included in the manuscript.

We expect that all researchers submitting to s-(PL)(DS)(Submissions)1(in whi)(A) h software is the central

Repositories













Tools and resources



Next steps