

Weekly Reading/Viewing

Please review the assigned material *before* class.

All items are posted on the Canvas site. Topics & readings are subject to change based on current events; changes will be announced via Canvas.

Week 1: Course Introduction

Week 2: File Formats, Naming Conventions, Data Storage, and Data Security.

Read:

1. Jan Čurn. 2014. How a bug in Dropbox permanently deleted my 8000 photos. [\[read online\]](#) [\[download pdf\]](#)
2. Panzarino, M. 2012. How Pixar's Toy Story 2 was deleted twice, once by technology and again for its own good. TNW. [\[read online\]](#) [\[\[download pdf\]\]](#)
3. Hart EM et al. (2016) Ten Simple Rules for Digital Data Storage. PLoS Comput Biol 12(10): e1005097. [\[read online\]](#) [\[download pdf\]](#)

Watch:

1. E. Bruna: [How To Name Your Files](#)
2. E. Bruna: [How to Organize Your Files](#)
3. E. Bruna: [How to Store and Backup Your Files](#)

Week 3: Data structure & format of data and datasets

Read

1. Tesi, W. 2020. An Outdated Version of Excel Led the U.K. to Undercount COVID-19 Cases. Slate. [\[read online\]](#) [\[download pdf\]](#)
2. Stolberg et al. 2020. CDC Test Counting Error Leaves Epidemiologists 'Really Baffled'. NY Times. [\[read online\]](#) [\[download pdf\]](#)
3. Broman, K. W., & Woo, K. H. (2018). Data organization in spreadsheets. The American Statistician, 72(1), 2-10. [\[read online\]](#) [\[download pdf\]](#)
4. Johnson, B. D., Dunlap, E., & Benoit, E. (2010). Organizing "mountains of words" for data analysis, both qualitative and quantitative. Substance Use & Misuse, 45(5), 648-70. [\[read online\]](#) [\[download pdf\]](#)

Week 4: Reproducible data (re)organization

Read:

1. Laskowski, 2020. What to do when you don't trust your data anymore. [\[read online\]](#) [\[download pdf\]](#)
2. Pennisi, E. 2020. Spider biologist denies suspicions of widespread data fraud in his animal personality research. Science. [\[read online\]](#) [\[download pdf\]](#)

3. Alston, J. M., and Rick, J. A.. 2020. A Beginner's Guide to Conducting Reproducible Research. Bull Ecol Soc Am 00(00):e01801. [\[read online\]](#) [\[download pdf\]](#)
4. Wilson G, Bryan J, Cranston K, Kitze J, Nederbragt L, Teal TK (2017) Good enough practices in scientific computing. PLoS Comput Biol 13(6): e1005510. [\[read online\]](#) [\[download pdf\]](#)

Watch:

1. [A Data Sharing and Management Snafu in 3 Short Acts, by the NYU Health Sciences Library](#)

Week 5: QA/QC 1: Data Entry & Validation.

Read:

1. Kamentz, A. 2018. The School Shootings that weren't. NPR [\[read online\]](#) [\[download pdf\]](#)
2. Lincoln, Matthew D. 2018. "Best Practices for Using Google Sheets in Your Data Project." [\[read online\]](#) [\[download pdf\]](#)

Week 6: QA/QC 2: Correction & Synthesis with Open Refine.

Read: *No readings this week.*

Week 7: QA/QC 3: Finding Errors with Data Visualization.

Read:

1. Rougier NP, Droettboom M, Bourne PE (2014) Ten Simple Rules for Better Figures. PLoS Comput Biol 10(9): e1003833. [\[read online\]](#) or [\[download pdf\]](#)

Watch:

1. Tommy McCall: [The simple genius of a good graphic](#)
2. R. Luke DuBois: [Insightful human portraits made from data](#)

Week 8: Documentation: Metadata, Codebooks; Data Sharing, Reuse, & Archives.

Read (Metadata):

1. Michener, W.K., et al . 1997. Non-geospatial metadata for the ecological sciences. Ecological Applications 7: 330–342. [\[read online\]](#) [\[download pdf\]](#)
2. Pp 446–450 in Bernard, H.R. and Bernard, H.R., 2013. Social research methods: Qualitative and quantitative approaches. Sage. [\[\[download pdf\]\]](#)

Read (Data Sharing):

1. Alexander, S.M., Jones, K., Bennett, N.J. et al. Qualitative data sharing and synthesis for sustainability science. Nat Sustain 3, 81–88 (2020). [\[read online\]](#) [\[download pdf\]](#)
2. Renaut, S. et al. 2018. Management, Archiving, and Sharing for Biologists and the Role of Research Institutions in the Technology-Oriented Age. *BioScience* 68(6)400–411 [\[read online\]](#) [\[download pdf\]](#)

3. Duke, Clifford S., and John H. Porter. 2013. "The ethics of data sharing and reuse in biology." *BioScience* 63(6): 483-489. [\[read online\]](#) [\[download pdf\]](#)
4. Tenopir, C., et al. (2015). Changes in data sharing and data reuse practices and perceptions among scientists worldwide. *PloS one*, 10(8), e0134826. [\[read online\]](#) [\[download pdf\]](#)
5. Mauthner, NS, & O. Parry O. 2009. Qualitative data preservation and sharing in the social sciences: On whose philosophical terms?. *Australian J of Social Issues* 44(3):291-307. [\[read online\]](#) [\[download pdf\]](#)

Week 9: Data Management Plans.

Read

1. Michener, W. K. (2015). Ten simple rules for creating a good data management plan. *PLoS Computational Biology*, 11(10), e1004525. [\[read online\]](#) or [\[download pdf\]](#)
2. Goodman A, et al. (2014) Ten Simple Rules for the Care and Feeding of Scientific Data. *PLoS Comput Biol* 10(4): e1003542. [\[read online\]](#) or [\[download pdf\]](#). this paper connects what we've done with what's coming up next

Watch

1. [Content of a DMP](#) (by RWTH Aachen University)

Week 10: Spring Break.

Week 11: Efficient data collection.

Read:

1. Redman, T. 2016. Bad Data Costs the U.S. \$3 Trillion Per Year. *Harvard Business Review*. [\[read online\]](#) [\[download pdf\]](#)

Watch

1. Atul Gawande's TED Talk: [The Importance & Value of the CHECK LIST](#).
2. Animated summary of [The Checklist Manifesto](#).
3. Jess Stratton's LinkedIn Learning Video Overview: "Use Google Forms to Create Surveys" (7 min.). Watch it [here](#); must be on UF computer or use UF proxy.
4. OPTIONAL: The more advanced, multi-video "Google Forms Essential Training Course". These short (1-3 min) videos explain each step in more detail. It's great, and only 39 min long from start to finish. Watch [here](#); must be on UF computer or use UF proxy.

Week 12: Transcription & Translation.

Read:

1. Bakker, Rebecca. "Transcription Tools and Software" (2017). *Works of the FIU Libraries*. 62. [\[read online\]](#) [\[download pdf\]](#)

2. Watch “The Text Wash team discusses text anonymization” [\[link\]](#); related blog post [\[link\]](#)

Watch

1. Dr. Jarek Kriukow: How to transcribe interviews - [Part 1: “naturalism” and “denaturalism”](#)
2. Dr. Jarek Kriukow: How to transcribe interviews - Part 2: [Which approach to use?](#)

Week 13: ‘Paperless’ data collection.

Read:

1. Aanensen DM, Huntley DM, Feil EJ, al-Own F, Spratt BG (2009) EpiCollect: Linking Smartphones to Web Applications for Epidemiology, Ecology and Community Data Collection. PLoS ONE 4(9): e6968. [\[read online\]](#) [\[download pdf\]](#)
2. Moylan, CA et al. 2013. Increasingly mobile: How new technologies can enhance qualitative research. Qualitative social work: research and practice, 14(1):36-47. [\[read online\]](#) [\[download pdf\]](#)
3. Teacher, Amber G. F. et al. Smartphones in ecology and evolution: a guide for the app-rehensive. Ecology and Evolution 3(16):5268– 5278 [\[read online\]](#) [\[download pdf\]](#)

Watch

1. E. Bruna: [Paperless Data Collection](#)

Great EpiCollect tutorials by UConn’s David Dickson

1. [Intro to EpiCollect 5](#)
2. [Creating an EpiCollect PROJECT](#)
3. [Creating a FORM for your Epicollect Project](#)
4. [Collecting data in EpiCollect with your phone or tablet](#)

Week 14: Automated data extraction.

Read:

1. Drinkwater, R. E., Cubey, R. W., & Haston, E. M. (2014). The use of Optical Character Recognition (OCR) in the digitization of herbarium specimen labels. PhytoKeys, (38), 15-30. [\[read online\]](#) [\[download pdf\]](#)
2. Joo, Jungseock and Zachary C. Steinert-Threlkeld. 2019. Image as data: Automated visual content analysis for social science. [\[read online\]](#) [\[download pdf\]](#)

Week 15: Legal and Ethical Issues.

Read:

1. de Koning M, Meyer B, Moors A, Pels P. 2019. Guidelines for anthropological research: Data management, ethics, and integrity. Ethnography. 20(2):170-174. [\[read online\]](#) [\[download pdf\]](#)

2. Katie Fortney, UC Office of Scholarly Communication. 2016. "Who owns your data?" [\[read online\]](#) [\[download pdf\]](#)
3. Goodluck, K. 2020. Indigenous data sovereignty shakes up research. High Country News. [\[read online\]](#) [\[\[download pdf\]\]](#)
4. Woodbury, R. B., Beans, J. A., Hiratsuka, V. Y., & Burke, W. (2019). Data Management in Health-Related Research Involving Indigenous Communities in the United States and Canada: A Scoping Review. *Frontiers in genetics*, 10, 942. [\[read online\]](#) [\[download pdf\]](#)

Watch

1. NIH: Statistical methods for protecting data confidentiality (<https://youtu.be/3vN-6A2zbsk>)
2. Patricia Jenkinson Video Seminar: [Ethics in Human Subject Research](#)
3. US Department of Health & Human Services: [Informed Consent Overview](#)
4. Australian Council for International Development (ACFID): [Principles and Guidelines for Ethical Research and Evaluation in Development](#)