ANT 298
Modern tools for data collection, management and analysis

Damien Caillaud
Department of Anthropology

August 29, 2016

This class will teach graduate students how to plan their data collection, data management system and data visualization early in their research. Using adequate, modern data management tools, researchers can avoid spending months entering and cleaning their data manually, learn how to create nice, interactive graphs, make simple web application and write R code and research manuscripts collaboratively.

Prerequisites: No previous knowledge of R or any other programming language is required.

Evaluation: each student will work on a personal project related to her/his topic of research and will develop **(i)** an electronic data collection tool adapted to her/his project, **(ii)** an SQL relational database and **(iii)** an `R shiny` web interface telling an interactive 'story" about her/his research (see a few examples here: `https://www.rstudio.com/products/shiny/shiny-user-showcase/`). Alternative datasets will be provided to students who do not have data to contribute.

Content

- structure of anthropological and biological data (*e.g.*, structure of ecological data, spatial archeological data, morphometric data, behavioral data . . .)

- electronic data collection tools: ODK (android), CyberTracker (android/windows mobile), Animal Observer (iOS)...

- relational database design, management and cloud storage using PostgreSQL

- data manipulation and visualization using **R**: querying a database, transforming data, using the basic **R** graph library and **R** GIS tools

- **R shiny**: a modern tool to develop web applications including D3 visuals and interactive maps

- version control using git, github, Overleaf to set up collaborative projects and write manuscripts collaboratively

- keeping up with technology: relevant literature, forums, websites . . .