Paradata for GSS 2021 (November 2022)

General Social Survey (GSS) paradata is collected variable information that captures contextual information about the setting in which the survey was completed. Historic paradata variables, sometimes called metadata, includes variables such as INTAGE and LNGTHINV, which captures the age of the interviewer conducting the survey and how much time was necessary to complete the interview, respectively. With the 2021 GSS conducted mostly through web surveys, it was necessary to introduce new paradata variables in order to continue to capture the context in which respondents completed the survey. This document defines and describes these newly introduce paradata variables, with special attention given to DEVICE and DEVICE_OS.

Important Note: The paradata is provided 'as-is". These paradata are direct extracts from data collection systems and they are defined based on different interactions the respondent may have had with the questionnaire. Some inconsistences are expected.

The full list of variables and a brief description of what they measure is included below:

DEVICE	Identifies the type of device with which the respondent completed the survey; for instance smartphone, tablet, desktop.
DEVICE_OS	Identifies the operating system of the device with which the respondent completed the survey, for example, Android, Linux, Windows, macOS.
META_USERAGENT	The browser respondent used
MODE_START	Survey mode respondents started with

Some important notes to understand the 2021 paradata are noted below:
- The new 2021 paradata variables are dynamic and should be associated with the last active session of the questionnaire. For example, a survey started on a desktop computer that breaks off and completes by phone will only report that it was completed on a smartphone. Similarly, a survey that was started on one device and with one web browser that was ultimately completed on another will only be recorded on that final device.
- META_USERAGENT values are reported to the GSS survey software when the survey is launched. If the survey is launched more than once, then the last browser and browser versions used to conduct the survey is reported.
- DEVICE and DEVICE_OS are most useful taken together as DEVICE_OS is inherently dependent on DEVICE. Users may find it helpful to generate a separate variable that combines these together. Given the variety of technological preferences by the public, there will be noise in the dataset for this recoded variable – one could always have a MacBook that runs a Windows OS, for example.

Below is a table describing how to understand different combinations of responses for DEVICE and DEVICE_OS.

DEVICE	DEVICE_OS	What this actually is
Smartphone	Android	Android-based smartphone, likely a Pixel or other Google-specific device that's running a minimally altered version of Android OS
Smartphone	Linux	Android-based smartphone, likely Samsung or other brand. It's possible that this may have captured some Android-based tablets as well.
Smartphone	Mac OS X	iPhone
Smartphone	Windows	An actual Windows-based smartphone, pretty small market share. Might want to recode this as 'other' for disclosure reasons.
Tablet	ChromeOS	Android-based tablet
Tablet	iOS	iPad, but an old device
Tablet	Mac OS X	iPad, newer device
Desktop	Linux	Laptop or desktop computer running Linux
Desktop	Chrome OS	Likely a Chromebook, a "light" version of a laptop, often used in schools
Desktop	Mac OS X	MacBook or iMac
Desktop	Windows	Windows-based laptop or desktop