

## Notes on Data Files Included in the CCRMA Virtual Archive

Andrew Nelson  
aj@ccrma.stanford.edu

- Files will download automatically upon clicking the web page link. All files are in Microsoft Excel format (except as noted)
- People and organizations are identified both by name and by ID numbers. The ID numbers were arbitrarily assigned, but are consistent across spreadsheets/datasets and thus permit cross-referencing and linking files.
- The Master ID List includes the ID numbers for each person and organization in the dataset. For organizations, it also includes location information (e.g., to facilitate geographic analyses)
- Data are current only through 2005 (and may have some lapses even leading up to 2005)
- These data are drawn primarily from the CCRMA Overviews, scans of which are also available on the CCRMA Virtual Archive. Additional data were drawn from individual CVs and other sources (e.g., personal communications with individual CCRMA-lites)
- Publication Citations data were compiled in 2009 and 2010 based on Google Scholar.
- Patent Citations data were compiled in 2009 and 2010 from the US Patent and Trademark Office.
- “CoreMemberComp” in the Compositions spreadsheet tracks whether or not the composer was a Stanford graduate student or faculty member (as opposed to a Visiting Scholar or some other affiliation)
- In the Patents spreadsheet, inventors who already have an ID are listed by that number. Inventors who otherwise do not appear in the CCRMA data are listed by their name.
- I’m working on updating all of these data and I welcome comments and corrections. That said, it may take me a bit of time. Thanks in advance for your patience.