Special Characters and Foreign Languages in REDCap

REDCap allows special characters and foreign languages for surveys or forms. However, importing or exporting data dictionaries or data with special characters can be difficult due to the default encoding in Microsoft Excel. This document will help you develop surveys and import and export data when special characters or foreign languages are used. Note that there are some things (e.g. "Next Page" and "Submit" buttons) that, for technical reasons, cannot be translated. Users can add descriptive text fields at the end of each page to translate these, and other buttons, as needed.

Building a Survey or Form

Online Designer

Any character that can be entered into a field label, field note, answer choice, etc. will be displayed properly in the REDCap system. Similarly, if survey participants use special characters in their answers, such as in a text or notes box, REDCap will display and store those characters properly.

Data Dictionary

If you prefer to build surveys or forms using the Data Dictionary upload, ensuring that special characters appear correctly can be trickier. Microsoft Excel saves all files with ANSI encoding, but UTF-8 encoding is necessary to display all special characters appropriately. Instead of Microsoft Excel, LibreOffice can be used to build the surveys or forms in the Data Dictionary. LibreOffice allows UTF-8 encoding. Some things to keep in mind when using the Data Dictionary to build projects with special characters:

- Always set encoding to UTF-8.
- Always save as a CSV file.

Projects with Multiple Languages

REDCap does not have the ability to set multiple labels and answer sets for a question, making conducting the same project in multiple languages tricky. However, there are some options and workarounds for a multi-language project.

1. Include the multiple languages in the same field label and choices. For example, a question asking how old someone is in both English and Spanish could say, "How old are you? / ¿Cuantos años tiene usted?"
2. Create multiple projects in different languages utilizing the same parent data dictionary after it has been translated into multiple languages. Merge exported data in the data analysis package of your choice after data collection is complete.
3. Create a language preference question at the start of a project or form so that a participant may select their language of choice. Then, set branching logic or use the survey queue to display the forms and questions in the participant's preferred language. This option requires that you have separate forms in each language, which means you will have many variables that must all be unique, even if they ask the same question in a different language. This isn't a barrier, but will add extra steps to the data cleaning and analysis phase.

Importing Data
Importing existing data that contains special characters also requires using LibreOffice or another software with UTF-8 encoding. Follow the instructions for importing existing data, using LibreOffice instead of Microsoft Excel. Again, make sure:

- Encoding is always set to UTF-8.
- The file is always saved as a CSV file.

Downloading CSV Files and PDFs

- Data Dictionaries or data with special characters will automatically export in a CSV file with UTF-8 encoding.
- Special characters do not properly export to PDF due to the default formatting of the PDF export module. Characters that appear normally in your browser or in an exported CSV file will appear as a jumbled mess when downloaded in a PDF. If you need to get a PDF of your instruments, the best workaround is to turn your codebook into a PDF.
    1. On the "Project Home" page, select the "Codebook" under "Quick Tasks."
    2. Select "Print page."
    3. Under "Destination," select "Save as PDF."
    4. Save the PDF by selecting "Save."

ILLINOIS
IHSI | Interdisciplinary Health
Sciences Institute