



Top: overview of a text-mining project. The center column shows the main steps, from data collection to the final analysis. The left column shows tools useful for the corresponding step. "new code" is written for a particular project. pubextract will be a Python package in which we plan to collect information extraction modules for specific aspects such as the number of participants. The right column shows the useful outputs produced by the project, and in parentheses suggestions of where they can be stored online. Bottom left: screenshot of labelbuddy annotating an article downloaded with pubget. Bottom right: screenshot of a small local web application displaying the results of a NeuroSynth-like meta-analysis, that is part of pubget's output when the "--fit\_neurosynth" option is used.