Figure 1: Overview of the Softcite Software extraction pipeline. Blue boxes represent the main processing components, and ovals the different data results.