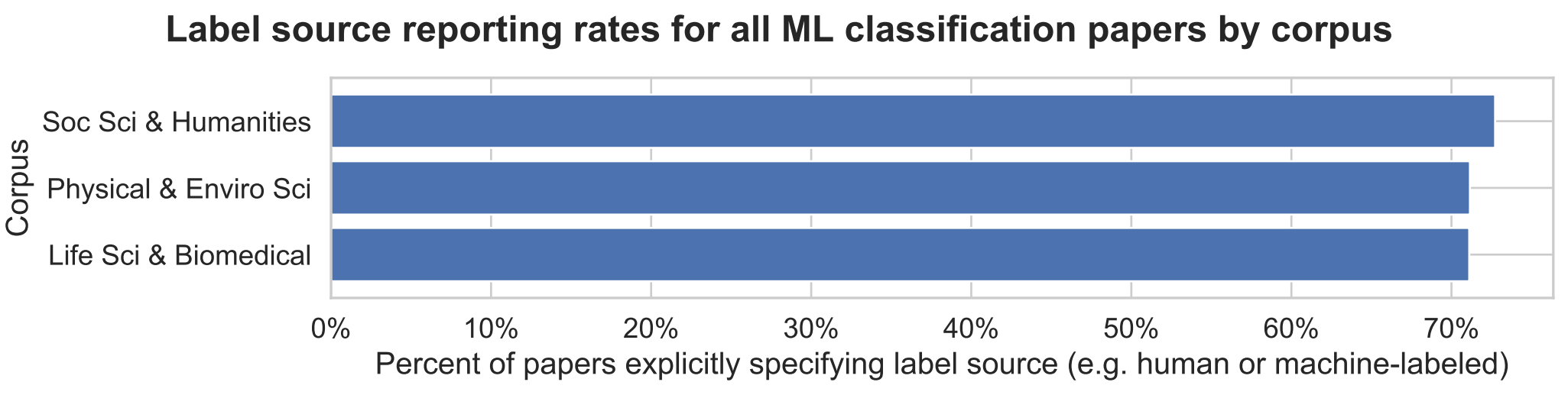
Label source reporting rates for all ML classification papers by corpus
Corpus
Soc Sci & Humanities
Physical & Enviro Sci
Life Sci & Biomedical
0%
10%
20%
30%
40%
50%
60%
70%
Percent of papers explicitly specifying label source (e.g. human or machine-labeled)