Named entities extracted from Old Books dataset with spaCy model
and compared with output from ground truth text. Each point represents
a noise/engine version (n=132) of the corpus. X axes cropped.

$$y = 0.76 - 0.0085\,x$$