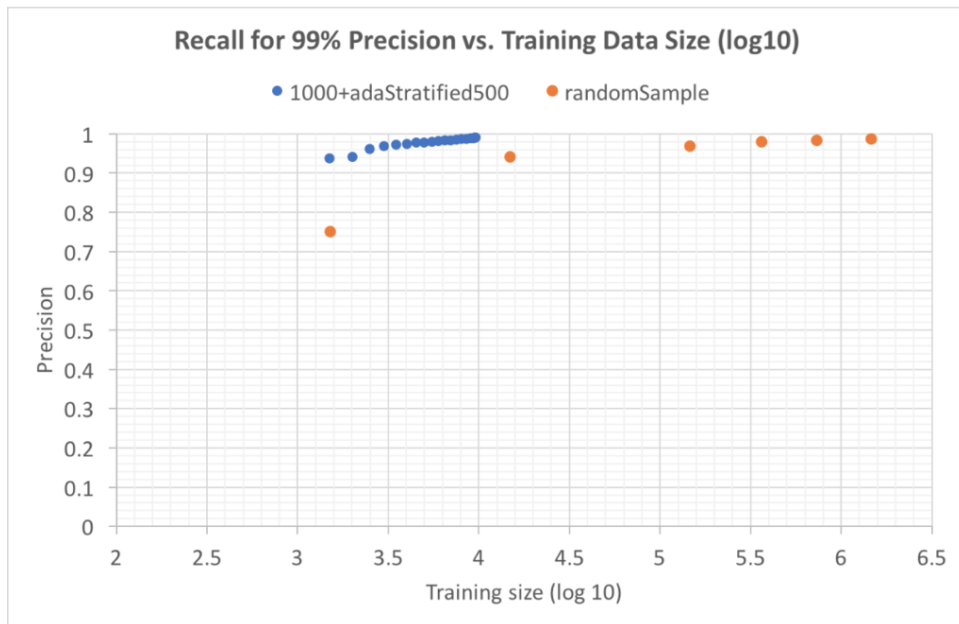


# Revisit Theme I. Which ML Model Works Best?

DI tasks	Hyperplanes (e.g., Log Reg)	Kernal (e.g., SVM)	Tree-based (e.g., Random forest)	Graphical models (e.g., CRF)	Logic programs (e.g, soft logic)	Neural networks (e.g., RNN)
Entity resolution	X	X	X		X	X
Data fusion	X			X		
DOM extraction	X					
Text extraction	X			X		X
Schema alignment	X			X		X

No single winner, although ensemble models and deep learning models show promising results.

# Revisit Theme II. Does Supervised Learning Apply to DI?



Active learning, semi-supervised methods, and weak supervision lead to dramatically more efficient solutions.