Evaluating a Set of Retrieved Docs: Precision and Recall

Action Doc	Retrieved	Not Retrieved
Relevant	Relevant Retrieved **a**	Relevant Rejected **b**
Not relevant	Irrelevant Retrieved **c**	Irrelevant Rejected **d**

$$\text{Precision} = \frac{a}{a+c}$$

$$\text{Recall} = \frac{a}{a+b}$$

Ideal results: Precision=Recall=1.0

In reality, high recall tends to be associated with low precision

Set can be defined by a cutoff (e.g., precision @ 10 docs)