QUALITY OF A SEARCH ENGINE

MEASURES
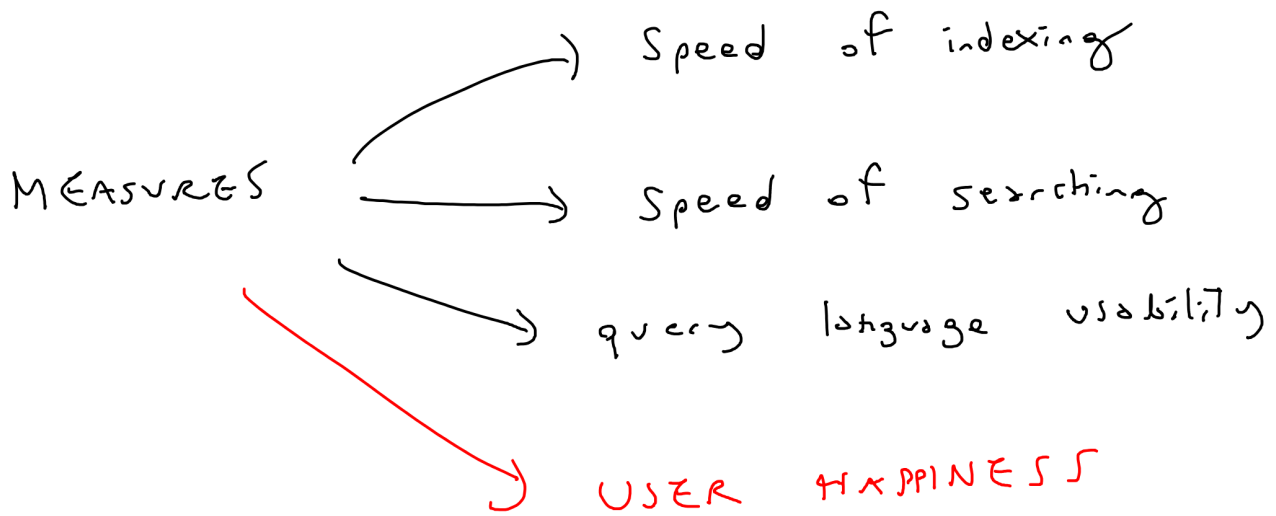- Speed of indexing
- Speed of searching
- query language usability
- USER HAPPINESS

The first two can be measured "in-house" (
while the last two requires a group of user
that evaluates the search engine

The user happiness depends mostly on the
content of the query result : are the documents
retrieved relevant or not? In which percentage?

NOTE = Also the UI/UX plays a role in
user experience

Precision and Recall metrics

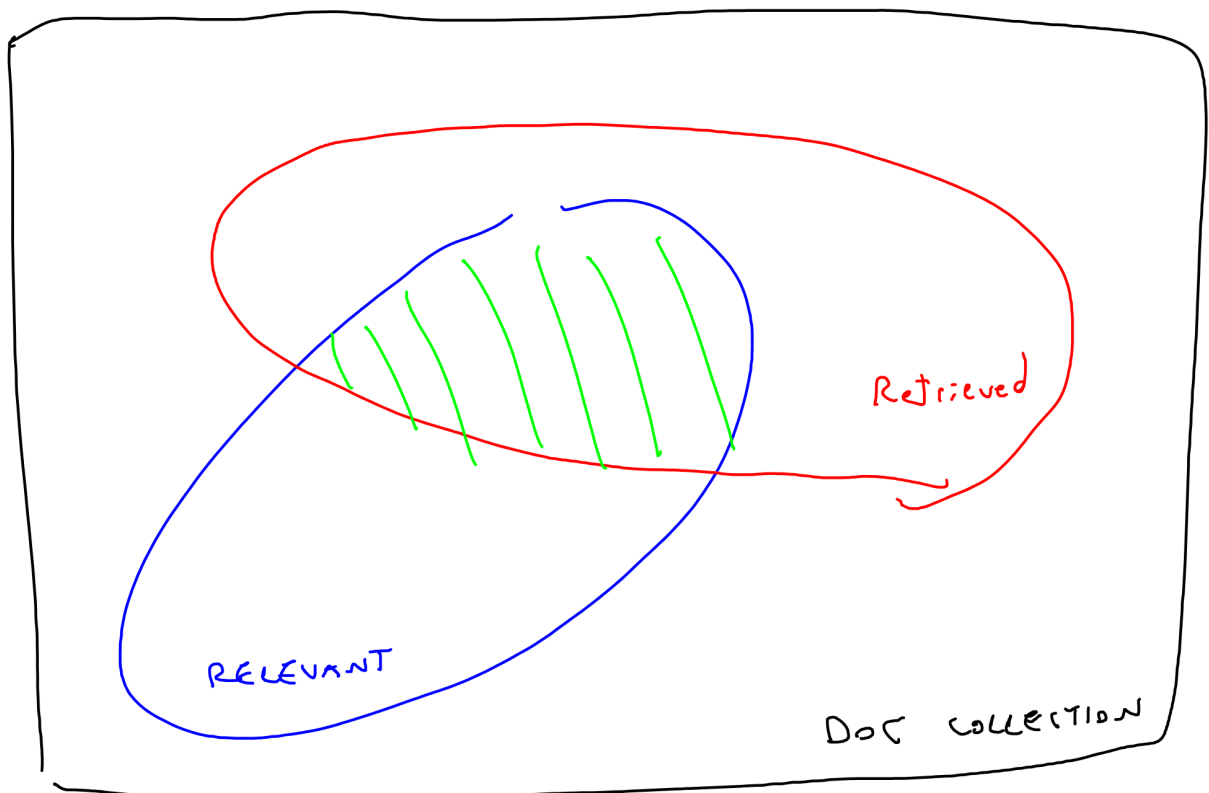Precision = fraction of retrieved docs that are relevant =

$$= \frac{\#(\text{Retrieved} \cap \text{Relevant})}{\#\text{Retrieved}} =$$

$$= \frac{TP}{TP + FP}$$
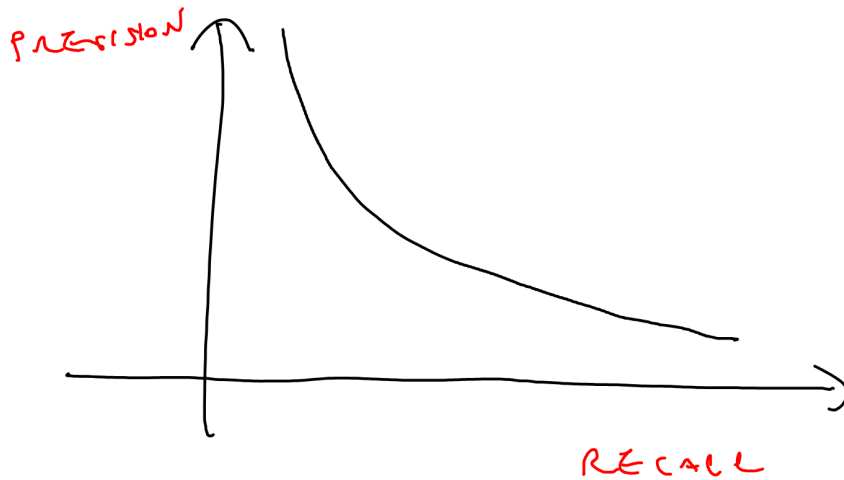
Recall = fraction of relevant docs that are retrieved =

$$= \frac{\#(\text{Relevant} \cap \text{Retrieved})}{\#\text{Relevant}} =$$

$$= \frac{TP}{TP + FN}$$

REMEMBER THAT

	Relevant	Not Relevant
Retrieved	TP	FP
Not Retrieved	FN	TN

PRECISION

RECALL

F - MEASURE

COMBINE PRECISION AND RECALL

$$F = \frac{1}{\alpha \cdot \frac{1}{P} + (1-\alpha) \cdot \frac{1}{R}}$$

if $\alpha > 1/2$ we give more weight to the precision, in respect to the recall.

For the balanced version $\alpha = 1/2$