# Recall vs. Precision vs. Summarization in RE for AI

**Daniel M. Berry**
**University of Waterloo**

# Tasks Requiring Intelligence

**We are talking about *tasks* requiring**

***real intelligence (RI)*, i.e.,**

**from a *human*.**

**The task is to find *correct answers***

**in a space of *answers*,**

**some *correct* and the rest *incorrect*.**

# Building an AI or LM

**We want to build an *artificial intelligence (AI)* that does the task.**

**This AI might be a *learned machine (LM)* which is the result of *machine learning (ML)*, whether it is taught, self-teaching, or both.**

# Specifying Requirements of AI

**How do we *specify the requirements* of the AI in a way that …**

**when we have an *implementaion* of the AI,**

**we can use the**

***requirements specification (RS)* of the AI**

**to decide whether the**

**implementation *meets***

**the AI's *requirements*?**

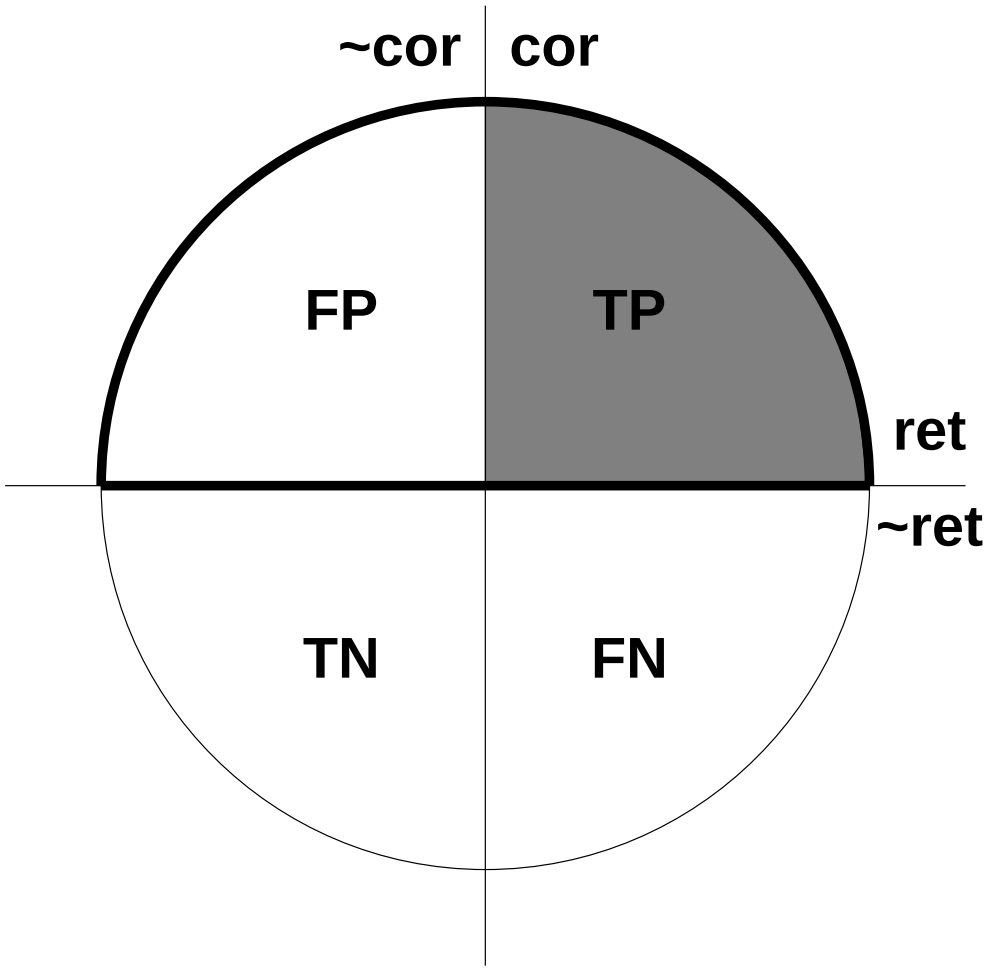
# Precision

***P* is the percentage of the tool-returned answers that are correct.**

$$P = \frac{|\ ret \cap cor\ |}{|\ ret\ |}$$

$$= \frac{|\ TP\ |}{|\ FP\ | + |\ TP\ |}$$

# Precision



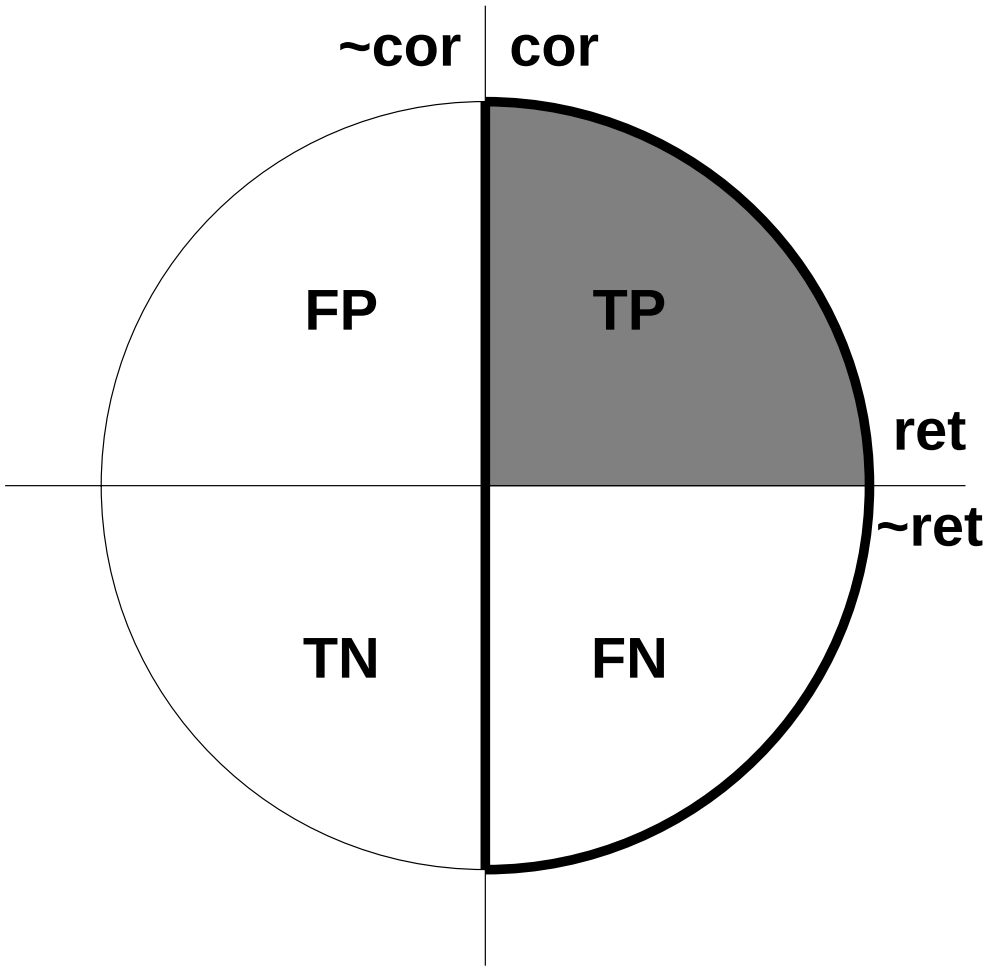
# Recall

***R* is the percentage of the correct answers that the tool returns.**

$$R = \frac{|\ ret \cap cor\ |}{|\ cor\ |}$$

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

# Recall



# *R* vs *P* Tradeoff

***P* and *R* can usually be traded off in an IR algorithm:**

- **increase *R* at the cost of decreasing *P*, or**
- **increase *P* at the cost of decreasing *R***

# Extremes of Tradeoff

**Extremes of this tradeoff are:**

1. **tool returns all possible answers, correct and incorrect: for**

   $R = 100\%$, $P = C$,

   where $C = \frac{\#\ \text{correctAnswers}}{\#\ \text{answers}}$

2. **tool returns only one answer, a correct one: for**

   $P = 100\%$, $R = \varepsilon$,

   where $\varepsilon = \frac{1}{\#\ \text{correctAnswers}}$

# Extremes are Useless

**Extremes are useless, because in either case, …**

**the entire task must be done manually on the original document in order to find *exactly* the correct answers.**

# 100% Recall Useless?

**Returning everything to get 100% *R* doesn't save any real work, because we still have to manually search the entire document.**

**This is why we are wary of claims of 100% *R* … Maybe it's a case of this phenomenon!**