

# Modélisation linguistique pour l'analyse automatique de textes

Aline ETIENNE  
acm.etienne@gmail.com

## Pourquoi et comment évaluer un système de TAL ?

- Impératif d'efficacité
- Besoin de mesures objectives
- Domaine de la recherche d'information : rappel, précision et f-mesure

### Rappel

$$R = \frac{\text{nb de bons résultats trouvés}}{\text{nb de résultats à trouver}}$$

### Précision

$$P = \frac{\text{nb de bons résultats trouvés}}{\text{nb de résultats trouvés}}$$

**/!\ Les valeurs de R, P et F-mesure sont toujours comprises entre 0 et 1**

F-mesure

$$F = 2 \cdot \frac{(\text{précision} \cdot \text{rappel})}{(\text{précision} + \text{rappel})}$$

## R, P et F-mesure pour évaluer un système

Le **petit** chat **roux** de ma **gentille** voisine Mireille dort sur le **vieux** paillason **gris**.  
X X ADJ ADJ X X ADJ X ADJ X X X ADJ X ADJ

➔ Reconnaissance des ADJ, R, P et F-mesure ?

$R = 4/5 = 0,8$  ou 80%

$P = 4/6 \simeq 0,67$  ou 67%

$F = 2 \times (R \times P) / (R + P) \simeq 2 \times (0,8 \times 0,67) / (0,8 + 0,67) \simeq 0,73$  ou 73%