

1h de vidéo HD sur Netflix : 1kg de O_2

taille des structures relation mots-pages :
combien d'octets pour 1 case ?
Nb de cases ?

score total de

$$s(d, r) = \alpha f(d, r) + \beta p(d)$$

page requête

score de fréquence TFIDF

pagerank

rééquilibrer f et p

$\approx 10^{-3}$ $\approx 10^{-6}$

?? $\alpha = 10^{-3}$ $\beta = 1 - \alpha$
 $\alpha f \approx 10^{-6}$ $\beta p \approx 10^{-6}$

1. calculer la moyenne de $f(d, r)$

2. ——— moyenne de $p(d)$

$\alpha = \frac{m_p}{m_f}$

$\beta = 1 - \alpha$

• tirer au hasard 10.000 requêtes de 1 à 5 mots
leur indice

• calculer $f(d, r)$ pour 1000 pages qui contiennent les mots de r .

→ on estime ainsi la moyenne $f(d, r)$