



Exemplar 1

pole
 scales
 plow
 harrow
 hoe
 .
 .
 .
 .
 .

Exemplar 2

pole
 rim
 scales
 plow
 harrow
 tool
 hoe
 .
 .
 .

$k = 2$

$SI_{\text{plow}} = \# \text{ shared genes w/in } k \text{ of plow}$

$SI_{\text{plow}} = 2 \text{ (scales, harrow)}$

$SI_{\text{doc pair}} = \frac{\text{sum}(SI_{\text{section}})}{\# \text{ shared genes}}$