

Evaluación en Recuperación de Información

Toda implementación debe ser validada para comprobar la eficacia del diseño. En este tema se revisarán las principales métricas aplicadas a los sistemas de información para comprobar su idoneidad.

El marco para validar los sistemas de recuperación de información más popular ha sido el paradigma Cranfield. Básicamente, se trata de un conjunto de consultas que se comparan con un conjunto predefinido de documentos, y manualmente se establece la relevancia entre ambos. Hasta la aparición de las competiciones TREC, este fue el principal marco de evaluación de sistemas de recuperación.

El paradigma Cranfield se basaba en pocos resultados, con una relevancia independiente del usuario y del tiempo. Las métricas aplicadas también estaban basadas inicialmente en un modelo simple. Estas métricas con el tiempo fueron teniendo en cuenta otros factores como: grado de relevancia (no tan solo de forma binaria), teniendo en cuenta el orden en que los documentos se recuperan y el grado en que la relevancia altera la relevancia conforme consulta documentos.

El paradigma Cranfield presenta el problema de su adecuación a Internet. En Internet la relevancia de cada uno de los documentos no es factible, ni siquiera la aplicación de algunas medidas clásicas como la recall.

Por último, se verán las medidas utilizadas en las competiciones TREC y posteriores marcos de evaluación.

