

Notas para la evaluación de la Práctica 2 (2ª Parte):

Motor de recuperación con Reconocimiento de Entidades y Expansión de Consultas

Ampliación del Motor de Recuperación Web

- Se evaluará el modelo conceptual de la aplicación, la adecuada implementación de los requisitos especificados para cada módulo, así como los resultados de la evaluación comparados con la parte primera.
- Recordad que el grountruth suministrado con la colección EIREX facilita ya los juicios de relevancia necesarios para la corrección de la práctica.

Módulo Base: Expansión de consultas

Autoevaluación

- Ha de tenerse en cuenta que la expansión de consultas puede aumentar el ruido en la recuperación. Los pesos asignados a los diferentes tipos de relaciones (sinonimia, jerarquía o relacionados), son cruciales para no obtener resultados no deseados.
- Hay que tener en cuenta que la expansión por determinado synset de Wordnet solo debería hacerse si el contexto de la acepción del término es el mismo, las técnicas utilizadas para lograr esto deberían ser implementadas en el programa, p.e. con LSI.
- Posibilidad de búsqueda por operadores: por campo, búsqueda por frase, por proximidad, con operadores booleanos, etc.
- Adaptación del modelo de recuperación para ajustar la relevancia en función de la expansión de términos.

Evaluación

Si el ruido no empeora y la accuracy mejora en comparación con las métricas de la parte 1, se considerará que la implementación es exitosa.





Autoevaluación

- Los principales problemas en la determinación de las entidades son debidos a los límites de la entidad y a los sinónimos (o alias) de la misma entidad. Implementar técnicas para lograr esto caracteriza una buena implementación.
- La aplicación adicional de heurísticas o expresiones regulares en el reconocimiento de personas.
- Reconocimiento de otros tipos de entidades (ej. fechas, localizaciones, etc.).
- La implementación de cualquier propuesta que suponga una mayor efectividad o eficiencia en la recuperación.
- La utilización de otros recursos adicionales (listados, ontologías, etc.), debidamente referenciados en la memoria y **sin realizar modificaciones sobre los mismos**.
- Al visualizar los resultados, resaltar las entidades que resultan más probables como respuesta.

Evaluación

Si las consultas que contienen un quién y que son identificadas como relevantes en el groundtruth son detectadas correctamente, la implementación será correcta.

