

Para [Peis,03] “a pesar de que el surgimiento de la Web ha supuesto un cambio radical en cuanto a la facilidad de acceso y difusión de la información electrónica, el gigantesco número de hiperdocumentos y su crecimiento exponencial, sumados a su falta de estructuración lógica, está dando lugar a graves problemas en la recuperación de información. Los actuales buscadores, basados en indización humana en el caso de los índices de búsqueda, e indización automática en el caso de los motores de búsqueda, son incapaces de ofrecer tasas de exhaustividad y precisión realmente aceptables en sus resultados. El problema de la precisión en la recuperación de información puede ser visto como consecuencia de la falta de significado o semántica que para los ordenadores tienen los documentos Web, en su amplia mayoría formateados mediante HTML, lenguaje de etiquetas o marcas que sólo especifican la forma de presentación de los contenidos”.

Así pues, si dotando de contenido semántico a los documentos se pudiera conseguir que las máquinas “comprendieran” el contenido de los mismos, los problemas antes mencionados de falta de exhaustividad y precisión de la recuperación de información se reducirían notablemente. En resumen, la Web semántica se basa en dos conceptos fundamentales:

- La descripción del significado que tienen los contenidos en el Web.
- La manipulación automática de esos significados.