Desde hace varios años, la cantidad de información disponible en la web crece de manera exponencial. Cada día se genera una gran cantidad de información que prácticamente de inmediato está disponible en la web. Los buscadores e indexadores recorren diariamente la web para encontrar toda esa información que se ha ido añadiendo y así, ponerla a disposición del usuario devolviéndola en los resultados de las búsquedas. Sin embargo, la cantidad de información es tan grande que debe ser preprocesada con anterioridad. Dado que el usuario que realiza una búsqueda de información solamente está interesado en la información relevante, no tiene sentido que los buscadores e indexadores procesen el resto de elementos de las páginas web. El procesado de elementos irrelevantes de páginas web supone un gasto de recursos innecesario, como por ejemplo espacio de almacenamiento, tiempo de procesamiento, uso de ancho de banda, etc. Se estima que entre el 40% y el 50% del contenido de las páginas web son elementos irrelevantes. Por eso, en los últimos 20 años se han desarrollado técnicas para la detección de elementos tanto relevantes como irrelevantes de páginas web. Este objetivo se puede abordar de diversas maneras, por lo que existen técnicas diametralmente distintas para afrontar el problema. Esta tesis se centra en el desarrollo de técnicas basadas en árboles DOM para la detección de diversas partes de las páginas web, como son el contenido principal, la plantilla, y el menú. La mayoría de técnicas existentes se centran en la detección de texto dentro del contenido principal de las páginas web, ya sea eliminando la plantilla de dichas páginas o detectando directamente el contenido principal. Las técnicas que proponemos no sólo son capaces de realizar la extracción de texto, sino que, bien por eliminación de plantilla o bien por detección del contenido principal, son capaces de aislar cualquier elemento relevante de las páginas web, como por ejemplo imágenes, animaciones, videos, etc. Dichas técnicas no sólo son útiles para buscadores y rastreadores, sino que también pueden ser útiles directamente para el usuario que navega por la web. Por ejemplo, en el caso de usuarios con diversidad funcional (como sería una ceguera) puede ser interesante la eliminación de elementos irrelevantes para facilitar la lectura (o escucha) de las páginas web. Para hacer las técnicas accesibles a todo el mundo, las hemos implementado como extensiones del navegador, y son compatibles con navegadores basados en Mozilla o en Chromium. Además, estas herramientas están públicamente disponibles para que cualquier persona interesada pueda acceder a ellas y continuar con la investigación si así lo deseara.