Protein  
   
Protein  
Proteins are large biomolecules and macromolecules that comprise one or more long chains of amino  
acid residues. Proteins perform a vast array of functions within organisms, including catalysing  
metabolic reactions, DNA replication, responding to stimuli, providing structure to cells and organisms,  
and transporting molecules from one location to another. Proteins differ from one another primarily in  
their sequence of amino acids, which is dictated by the nucleotide sequence of

## Wikipedia Excerpt: Document classification

Document classification or document categorization is a problem in library science, information science and computer science. The task is to assign a document to one or more classes or categories. This may be done "manually" (or "intellectually") or algorithmically. The intellectual classification of documents has mostly been the province of library science, while the algorithmic classification of documents is mainly in information science and computer science. The problems are overlapping, however, and there is therefore interdisciplinary research on document classification.  
The documents to be classified may be texts, images, music, etc. Each kind of document possesses its special classification problems. When not otherwise specified, text classification is implied.  
Documents may be classified according to their subjects or according to other attributes (such as document type, author, printing year etc.). In the rest of this article only subject classification is considered. There ar