



## Full-Text (Substring) Indexes in External Memory

By Marina Barsky

Morgan & Claypool. Paperback. Condition: New. 92 pages. Dimensions: 9.1in. x 7.4in. x 0.3in. Nowadays, textual databases are among the most rapidly growing collections of data. Some of these collections contain a new type of data that differs from classical numerical or textual data. These are long sequences of symbols, not divided into well-separated small tokens (words). The most prominent among such collections are databases of biological sequences, which are experiencing today an unprecedented growth rate. Starting in 2008, the 1000 Genomes Project has been launched with the ultimate goal of collecting sequences of additional 1,500 Human genomes, 500 each of European, African, and East Asian origin. This will produce an extensive catalog of Human genetic variations. The size of just the raw sequences in this catalog would be about 5 terabytes. Querying strings without well-separated tokens poses a different set of challenges, typically addressed by building full-text indexes, which provide effective structures to index all the substrings of the given strings. Since full-text indexes occupy more space than the raw data, it is often necessary to use disk space for their construction. However, until recently, the construction of full-text indexes in secondary storage was considered impractical due to excessive...



[READ ONLINE](#)  
[ 3.29 MB ]

### Reviews

*A must buy book if you need to adding benefit. it absolutely was writtern very properly and valuable. I found out this book from my i and dad advised this ebook to find out.*

-- **Amanda Larkin**

*A fresh electronic book with a new perspective. It is one of the most remarkable book we have go through. Your daily life period will likely be transform the instant you full reading this article pdf.*

-- **Katrine Kohler DVM**