How do Internet search engines work?

-A. Dharia, Houston

Javed Mostafa, Victor H. Yngve Associate Professor of Information Research Science and director of the Laboratory of Applied Informatics, Indiana University, offers this answer:

Publicly available Web services—such as Google, InfoSeek, Northernlight and AltaVista—employ various techniques to speed up and refine their searches. The three most common methods are known as preprocessing the data, "smart" representation and prioritizing the results.

One way to save search time is to match the Web user's query against an index file of preprocessed data stored in one location, instead of sorting through millions of Web sites. To update the preprocessed data, software called a crawler is sent

periodically by the database to collect Web pages. A different program parses the retrieved pages to extract search words. These words are stored, along with the links to the corresponding pages, in the index file. New user queries are then matched against this index file.

Smart representation refers

to selecting an index structure that minimizes search time. Data are far more efficiently organized in a "tree" than in a sequential list. In an index tree, the search starts at the "top," or root node. For search terms that start with letters that are earlier in the alphabet than the node word, the search proceeds down a "left" branch; for later letters, "right." At each subsequent node there are further branches to try, until the search term is either found or established as not being on the tree.

The URLs, or links, produced as a result of such searches are usually numerous. But because of ambiguities of language (consider "window blind" versus "blind ambition"), the resulting links would generally not be equally relevant. To glean the most pertinent records, the search algorithm applies ranking strategies. A common method, known as term-frequencyinverse document-frequency, determines relative weights for words to signify their importance in individual documents; the weights are based on the distribution of the words and the frequency with which they occur. Words that occur very often (such as "or," "to" and "with") and that appear in many documents have substantially less weight than do words that appear in relatively few documents and are semantically more relevant.

Link analysis is another weighting strategy. This technique considers the nature of each page—namely, if it is an "authority" (a number of other pages point to it) or a "hub" (it points to a number of other pages). The highly successful Google search engine uses this method to polish searches.

What is quicksand?

—S. Yamasaki, Brussels, Belgium

Darrel G.F. Long, a sedimentologist in the department of earth sciences, Laurentian University in Sudbury, Ontario, explains:

Quicksand is a mixture of sand and water or of sand and air; it looks solid but becomes unstable when it is disturbed by any additional stress. Grains frequently are elongated rather than spherical, so loose packing can produce a configuration in which the spaces between the granules, or voids, filled with air or water make up 30 to 70 percent of the total volume. This arrangement is similar to a house of cards, in which the space between the cards is significantly greater than the space occupied by the cards. In quicksand, the sand collapses, or becomes "quick," when force from loading, vibration or the upward migration of water overcomes the friction holding the particles in place. In normal sand, in contrast, tight packing forms a rigid mass, with voids making up only about 25 to 30 percent of the volume.

Most quicksand occurs in settings where there are natural springs, such as at the base of alluvial fans (cone-shaped bodies of sand and gravel formed by rivers flowing from mountains), along riverbanks or on beaches at low tide. Quicksand does appear in deserts, on the loosely packed, downwind sides of dunes, but this is rare. And the amount of sinking is limited to a few centimeters, because once the air in the voids is expelled, the grains nestle too close together to allow further compaction.

For a complete text of these and other answers from scientists in diverse fields, visit www.sciam.com/askexpert