**Description**

# What is a Search Algorithm?

In layman's terms, an algorithm is a set of coding instructions or rules for a computer to follow in completing a calculation or problem-solving operation. A search algorithm, as the name implies, is written to find items with specific properties within an increasingly large database of items.

A *linear* search looks at each item in sequence until it finds the items it was written to look for and then stops. As the size of the database increases, the algorithms get more sophisticated. *Binary* searches look for records tagged by a numerical key or field such as your driver's license number or social security number. Depending on how the data has been structured, a "tree" search will start at a root item and build results in a tree pattern. At the top of the list is a genetic algorithm that searches for an "optimum solution" based on available data by testing strings of data and discarding the worst ones in multiple iterations until the "best" is left.

# A Vital Research Tool

The most familiar search algorithm is probably Google's, and the non-threatening names given to its frequent updates—Panda, Penguin, Pigeon—help to minimize the nerd factor for the non-techies that use it multiple times every day. However, when a simple search for the term "penguin" on Google delivers 165 million results in only 0.35 seconds, the size of the databases being searched and the dire need for an accurate tool to search them becomes evident.

In Academia, there are estimated to be over 110 million scholarly documents available in English on the Internet, with the majority of them containing topic-specific terminology. Google Scholar has already started to make inroads into the vast trove of knowledge, but critics are already complaining that the algorithm makes no effort to distinguish between refereed and non-refereed sources. As a result, questionable research work gets listed in the same results as peer-reviewed research of established quality.

# Failing to Keep-Up

If you subscribe to the old axiom that "a rising tide raises all boats," you would think that the rapid advances in digital technology that are facilitating the construction of mega databases with petabytes and yottabytes of data, are also being matched with the coding knowledge to expertly navigate through all that data. That is not the case. In the early days of computing, the old adage of *Garbage In Garbage Out* (GIGO) reminded programmers that the computer could only function as efficiently as the code written. In other words, stupid code made the computer do stupid things. The sheer volume of data, with which we are now working has escalated that problem to an exponential level. In academic terms, a poorly structured search string in a database with tens, if not hundreds, of thousands of entries, could produce results that could totally derail your research before the literature is even complete.



Copyright: Enago Academy under the CC-BY-NC-SA 4.0 license

# From Searching to Predicting

We are already seeing multiple examples of questionable use of powerful algorithms. Edward Snowden's revelations of the complex algorithms developed by the National Security Administration (NSA) to track suspected terrorists, demonstrated just how much information was available on American citizens, and how detailed the developed profiles of those citizens could be, based on known information and predicted data built on statistical calculations of relevant averages.

For academic research, the sheer volume of data is now attracting increasingly complex meta-analyses of specific areas of research that are, in turn, driving the development of predictive software to help institutions and corporations to decide which specific areas of research will have the greatest chance of success. Critics argue that we're one step away from treating research like a Las Vegas "system," but one thing is clear, unless personnel are trained in the correct use of these powerful algorithms, there's going to be even more questionable data floating around to waste increasingly scarce research funding.

**Category**

1. Publishing Research
2. Submitting Manuscripts

**Date Created**
2016/03/25
**Author**
daveishan



Copyright: Enago Academy under the CC-BY-NC-SA 4.0 license