

Finger Printing

Human fingerprints are defined by tiny ridges, spirals and valley patterns on the tip of each finger. They are unique: no two people have the same ones.

Similarly, a fingerprinting algorithm in computer science is one that maps large data (such as documents or images) to a much shorter sequence of bytes. Such a sequence may be called the data's fingerprint. In essence, large data is more tersely represented by its fingerprint. This fingerprint uniquely identifies the original data, just as human fingerprints uniquely identify individual people. Fingerprinting can be applied to different types of inputs: documents, images, audio, video, or any arbitrary data.

While fingerprints may identify the original data, they're not a compression technique. This means that original data cannot be derived from its fingerprint. Fingerprints help in quickly checking the integrity of data or retrieving documents from large filesystems.

