Metadata 101

There are many different forms of digital evidence. For purposes of trial, this evidence can be broken down into three broad categories: (1) digital evidence created by a machine; (2) digital evidence created by a human; and (3) digital evidence recovered by an expert (often a police detective).

Evidence created by a machine is self authenticating for legal purposes. The foundational requirements are merely a showing that the machine is working properly at the time the record was made. The greater challenge is tying the computer data to a specific person. This requires various types of circumstantial evidence that places the defendant at the keyboard at the time the record was created. Recovering data of a personal nature (i.e. password protected data with access dates and times at or near in time to the target evidence can be very effective.

This leads to Digital evidence created by a human user interfacing with the machine, typically the person that discovered something wrong (i.e. computer technician, Network IT professional, co-worker, spouse, etc.) who has somehow created a digital document or account/ summary of their discovery. In this situation, they are presented as a fact witness describing the circumstance leading up to the discovery of the evidence at issue. Whatever digital evidence they created is hearsay unless presented through that witness or falling under a recognized hearsay exception.

The final type of digital evidence is that which the expert recovers from the machine. This requires two witnesses; the examiner who recovered the data and the witness who can link this data to the defendant or person of interest. In some jurisdictions, the experts and the detectives are one in the same; in others the standard practice is to separate the two functions.

Emails are the most common type of digital evidence sought to be introduced. The ubiquitous combination of text and images can provide very compelling evidence to a jury whose attention span is limited. Emails are a hybrid of two of the three types of digital evidence described above. The recipient header information is generated by the machine, but the re: line and text is generated by the human user. Most, if not all, emails can be recovered by an expert.

It is important to note that when dealing with a computer hard drive, or other type of digital information storage device, the data is nothing more than 1's and 0's arranged in a pattern on the disk (sometimes called the platter). There are two types of "space" on a computer hard drive: (1) allocated and (2) unallocated. Allocated space means that the computer sees this information stored there and you can access it. Unallocated space means the free space, or space available for storing digital information.

What happens when you create a document for instance, is that the computer codes that digital information into 1's and 0's and then finds a place in the disk's unallocated space to now store that information for later use/ retrieval. Once saved, that information is now in the disk's allocated space.

The computer hard drive is like a paper back book, when you save something, you are creating a "bookmark" or entry in the book's table of contents. When you later delete the document, all you have actually done is removed the "bookmark" or the page number from the table of contents. The digital item remains on the computer's disk, but has now been designated as "available space" or back to unallocated space, allowing some other digital information to be written over it.

The only true way to "delete" information is to "wipe the drive" which means to write random 1's and 0's or just 1's or just 0's over the disk many times. Government standards for wiping are 7 "passes" over the disk. CIA standards are like 50 "passes" over the disk. If you ever encounter a disk that has been "wiped" then there's usually a strong case of evidence tampering that can be charged. At least I also did.

Meta data is basically the name of automatically generated and stored data from a machine such as email platforms, digital cameras, Smartphone, etc. This will give you specific information about the device such as make, model, serial number, lens exposure, date, time, etc. Very helpful if you can get it. Typically only recoverable by way of special computer forensic tools, however, it will only be as reliable as the human how performed the initial setup. For instance, in most cameras, when the batteries run out, the user will have to reset the date and time otherwise it defaults back to 12:00 am January 1, 2000, or some other da te and time. While other information may still be accurate, the date and time are usually the most sought after.

There are a number of commercial grade programs available that "Scrub" the meta data off of a document or other digital information, so that important internal information does not get released to an unauthorized third party. However, even when "scrubbed" the data may be recovered forensically from the original computer or device that sent the information.