6 DOs & DON'Ts:

For Data Archiving

> DO. Clearly differentiate between your archive and your backup solutions.

During the past couple of years, some vendors have attempted to blur the line between backups and archives. Even so, backups and archives serve two distinctly different purposes and should be treated as such.

Backups are created as a data recovery mechanism. In contrast, archives are designed for long-term storage of data and are not typically treated as a disaster recovery mechanism.

Because of this distinction — and the fact that archived data should never be modified — some healthcare organizations do not back up or replicate their archives. However, archives can be prone to failures resulting in data loss, just as any other type of IT system can. Although traditional nightly backups probably aren't appropriate for archives, you do need a redundant copy of archived data. For example, an organization that keeps its archives on storage area network (SAN) storage might use storage replication as a way of ensuring that a redundant copy exists, and is automatically kept current or can be replicated to the cloud to ensure that it is available in the event of a disaster.

> DON'T. Configure archive storage in a way that could potentially become a single point of failure.

All too often, an organization's archives are treated as something of an afterthought. Data retention is an operational requirement, but because archives are filled with data that is rarely accessed, there might be pressure to use the least expensive storage solution possible.

There is nothing wrong with using low-cost hardware for storing archived data, but the archives must not be constructed in a way that could result in a single point of failure. If an organization's archive system were to fail, a huge volume of data could be lost — and HIPAA regulations stipulate severe penalties for this type of data loss.

Even if a component failure didn't cause any actual data loss, it would result in downtime for the archive system. Archive contents might be rarely accessed, but when someone does access the archives, there is usually an important reason for it. As such, it is important to construct the archive system in a way that prevents any component from becoming a single point of failure. Doing so can help prevent data loss and outages.

TechTarget Custom Media

IRON MOUNTAIN®

> DO. Establish clear policies defining what data should be stored in which tier of storage.

The volume of data within a healthcare provider's archives can be huge, and there are costs associated with storing vast quantities of data. Organizations typically attempt to curb this cost by using multiple storage tiers. For example, archives might be spread among SAN storage, cloud storage and removable media such as tape, with each medium acting as a separate storage tier.

It is important to have policies in place that stipulate what each storage tier is used for. These policies are usually based around the age of the data, data type or some combination of the two. For example, a healthcare provider may wish to store patient health data that is less than 5 years old on spinning disk so that it is readily accessible, while moving older data to tape or the cloud. This approach can help the organization keep the most current data readily accessible, while also reducing storage costs by placing the oldest data onto less expensive media.

> DON'T. Treat all of your data the same with regards to archival.

Just as it is important to have policies in place that dictate how the various tiers of archival storage are to be used, it is also important to realize that it is a mistake to treat all of your data equally with regard to archiving. Data archiving requirements tend to vary widely based on data type.

These variances are often based on retention requirements. For example, a healthcare provider might have internal operational requirements that require it to retain patient data for 10 years after the patient has died. That same organization might need to keep financial data for only seven years.

It is also important not to overlook the fact that some data may not need to be archived. For example, a healthcare organization probably would not need long-term archiving of marketing materials or PowerPoint presentations unless there were operational requirements to do so.

> DO. Take steps to automate data lifecycle management.

Regardless of whether a healthcare provider is large or small, it is unrealistic to expect that the data archiving process can be done manually. Manual archiving is impractical for a number of reasons.

For starters, like any other manual process, manually archiving data is a labor-intensive process. Although this process might at first seem manageable, history has shown that data grows exponentially over time, and healthcare data growth rates today are nearing

40%. Therefore, what might be a somewhat manageable process today will likely become unwieldy and out of control later on. Even if the process were to somehow remain manageable for an indefinite period of time, manually archiving data is a poor use of your IT staff's time, especially when you consider that there are automated data lifecycle management products readily available.

Another reason why you should automate data lifecycle management is that it is difficult to perform manual archiving on a consistent basis. Manually performing data lifecycle management probably isn't one of those tasks that an organization performs on a daily basis. As such, it is easy for the IT staff to forget to periodically manage the data.

A third reason why manual data lifecycle management is a bad idea is because the process is prone to human error. It is all too easy for data to accidentally be moved to the wrong storage tier or to be prematurely purged. In the case of patient health data, there can be severe consequences for this type of data mishandling. An upfront investment in data lifecycle automation software can save a substantial amount of money in the future in the form of labor costs and, potentially, in the avoidance of fines from the U.S. Department of Health and Human Services.

> DON'T. Underestimate the importance of security — especially when archiving data to the cloud.

Even if your archives consist entirely of very old data, the content is sensitive. Organizations do not archive data unless it is potentially useful in the future or is needed to address compliance around long-term retention requirements. Archives may contain financial data, protected electronic health information, personally identifiable information about the organization's employees, or just about anything else that can be imagined. Given the sensitivity of the archived data, it is critically important to make sure that the archives remain secure. This is especially true when archiving data to a cloud storage provider, since the cloud storage is outside of your direct control.

The Omnibus rule raises the stakes even further. Under the rule, cloud storage providers are treated as business associates. Furthermore, a covered entity shares the responsibility for security breaches with its business associates. This is why it is so critically important for healthcare providers to evaluate a cloud storage provider's security and to ensure that the provider is fully HIPAA-compliant.

> DO. Assess the benefits and risks of various storage solutions.

There are a number of different options for storing archived data, and there are advantages and disadvantages to each type of storage. As such, it is important to examine the pros and cons of the available storage solutions before committing to an archive method. For example, tape has the advantage of being a highly reliable storage medium with a very low cost per gigabyte of storage. On the other hand, data that is stored on tape is less accessible (the tape must be mounted before the data can be read). Furthermore, the portable nature of tape makes it possible for vast quantities of archived data to walk out the door, so it is important to ensure your tapes are tracked and stored using a secure chain of custody.

On-premises online storage is another popular option. This option uses spinning disks to store archived data, usually on a SAN or network-attached storage device. The main advantages to using this type of storage are that the data is easily accessible and that fault-tolerant technologies can be used to safeguard against disk failures. Conversely, this type of storage tends to be significantly more expensive than removable media. Although on-premises online storage can be scaled, doing so consumes floor space and increases electric and cooling costs. Furthermore, the storage will eventually become obsolete, thereby leading to hardware replacement and data migration costs.

Cloud-based storage is often thought of as a happy medium between on-premises disk storage and tape storage. Cloud storage is generally less expensive than on-premises disk storage, and storage capacity ceases to be an issue. However, the fact that your data is being stored outside of your direct control means that security and the cloud storage provider's reputation are of critical importance.

> DON'T. Try to go it alone.

Finally, don't go it alone when implementing a data archiving system. Healthcare providers are subject to very strict data retention requirements, and there can be severe penalties for failing to adhere to those requirements. Given the exceedingly high stakes surrounding data lifecycle management in healthcare organizations, it is a good idea to enlist the help of an organization that focuses on helping healthcare IT to properly implement data archiving and data lifecycle management solutions.