



Making Office 365 a One-Stop-Shop for Email Records Compliance

With recent enhancements to its information management and eDiscovery capabilities, Microsoft Office 365 is rapidly reaching the point where organizations no longer need to manage email records on-premises or in third-party services in order to meet their compliance needs.

This whitepaper offers organizations planning a migration to the Cloud a 'considerations roadmap' for migrating their legacy email records.

Why should I read this paper?

In this white paper we will provide guidance on the critical aspects of sustaining the integrity and value of your legacy email records as they are migrated to Microsoft Office 365.

Key insights include:

	Page
• 10 things to consider as you migrate legacy email records to Office 365	7
• Moving journals into the new Office 365 model	14
• How Microsoft replaces the journal in Office 365	15
• Enhanced eDiscovery in Office 365	18

Additionally this paper is positioned as an **advisory piece for all key stakeholders** that should be involved in a migration project. That is, not just the IT team, but also the legal department, records managers, business leaders and advocates acting on behalf of users.



Over the past few years Microsoft's Office 365 has been attracting increasingly larger enterprises with the -promise of predictable costs, easier scalability, built-in service resilience (backed up by 99.9% uptimes) and a platform that caters for an increasingly mobile and collaborative workforce.

However, what about those organizations that have been holding back with a fully in-house Exchange environment, or hedging their bets with a hybrid setup?

Another significant hurdle that many have had to contend with is whether Office 365 could fully meet their data security and compliance needs.

But this barrier is now falling away:

Along with getting security and privacy certifications relevant to virtually every vertical market segment, inclusive of regulation heavyweights like [Government and Health](#), and advanced data loss prevention controls, Microsoft has been busy evolving Office 365 to the meet the information governance and eDiscovery demands of enterprise email customers.

Services added over the last few years include:

- Immutable preservation of emails to meet retention and eDiscovery demands.
- Indefinite storage of 'inactive' user mailboxes.
- Preservation of vital compliance metadata, such as all staff members named in email distribution lists.
- The [Compliance Center](#) – a dashboard designed to enable non-technical staff to manage compliance-related activities.

The [recent acquisition of Equivio](#) has added an Azure-based Early Case Assessment (ECA) tool which will further appeal to enterprises seeking to reduce the costs, time and risks in preparing for litigation 'in house'.

Features like this will speak to any IT decision maker with compliance or corporate governance under their set of responsibilities.

Is Microsoft Office 365 now a one-stop shop if you have a compliance or corporate governance remit?

For many enterprises, Office 365 could be the golden bullet. However to realize its full potential, they'll need to **expertly move** what could amount to many terabytes-worth of legacy email records into it.

Without this consolidation, enterprises will bear the overheads, costs and risks of maintaining the accessibility of data, managing its lifecycle and performing eDiscovery using multiple different interfaces across disjointed repositories that include:

- Office 365
- On-premises Exchange servers & Journal mailboxes
- Third-party email archives
- PST files
- Hosted email services

Critically, where the process of consolidating (migrating) such data into Office 365 is executed without a firm handle on the compliance issues and pitfalls at stake, the results could be rendered unreliable and unacceptable when they are needed most: when trying to win or defend a court case; uncover the truth behind an internal HR dispute or sort out a multi-million business transaction 'gone wrong'.

Likewise, a compliance-led migration that is not people-centric in its approach will result in a significant productivity impact.

Those that hold the purse strings are excited by the financial benefits of not having to source additional services and maintain separate systems when they can get everything in one place...

FOREWARD by STEVE GOODMAN, MVP

The Evolution of Office 365, Information Governance & eDiscovery

Microsoft Exchange has an interesting relationship with compliance features, and over the course of its history has evolved significantly - from lacking any significant features to providing the full complement of features most organizations need in a modern business environment.

1998

Back in 1998, Exchange 5.5. Service Pack 1 introduced the ability to add a journal recipient to the Exchange system. This allowed an administrator to drop off a copy of every message sent and received within the organization to one or more specific mailboxes. Due to the size requirements, long-term use of journal mailboxes was and remains impractical for many organizations, so many use a third-party solution like Symantec Enterprise Vault to extract and keep this data.

2009

Exchange Server has moved on significantly and made the biggest leap with the release of Exchange 2010. Retention Policies allowed organizations to remove content after it became no longer relevant. Litigation hold functionality allowed content to be kept intact and unchanged within a mailbox even if a user deleted or changed it, and then discovered when required.

Fantastic as those features were, many on-premises customers felt a third-party solution was still needed both to offload the Exchange storage requirements and backup requirements, combined with the risk that functionality like litigation hold was a version-one technology. Additionally organizations wanting to take on the mantle of eDiscovery in-house would look to specialized vendors.

2011

Enter Office 365 - Office 365 was released to businesses in 2011 as the successor to Microsoft's less popular Exchange 2007 service, BPOS (Business Productivity Online Suite) and Exchange Online (evolved from Exchange Labs and Live@EDU).

Upon launch, Office 365 used Exchange 2010 for the mail service and it represented a major shift in the way that Microsoft designed and built server software. Rather than architect the software for a customer's on-premises datacentres first and then retro-fit for the cloud, Exchange 2010 was developed for the cloud first. **This brought about the biggest jump in functionality seen in a version of Exchange as Microsoft could not rely on third-party products and needed to provide very large mailboxes, very reliably at low cost and without backups.**

Early customers of Office 365 faced some challenges with compliance and this held organizations with such requirements back from choosing the service. Early adopters would often retain an on-premises system or use a third-party cloud service to maintain an email journal to meet compliance needs. *Even though Office 365 now provides a 50GB primary mailbox and in some plans, effectively unlimited archives, journal mailboxes are not allowed within the service terms.*

2013

The Wave 15 major upgrade to Office 365 coincided with the launch of Exchange Server 2013. This upgrade expanded upon the features released with Exchange 2013, including In-Place Hold.

In-Place Hold changed the perception of Office 365 for many organizations hesitant about functionality offered. Similar to the litigation hold feature, where any deleted or modified items are kept within a read-only portion of each mailbox, In-Place Hold allows organizations to create a set of policies that match business requirements, such as keeping all mail for 6 years for a particular group of users.

The improved hold functionality is complemented by re-tooled eDiscovery tools; first built into the Exchange admin centre to cover discovery and extraction of email-related content, and across the Office 365 suite via the **eDiscovery Center** which allows full case management and data export, including relevant XML metadata in a format that complies with the Electronic Discovery Reference Model (EDRM) framework. Office 365 discovery tools allow rich searches using advanced search terms and with recent acquisitions by Microsoft and integration of technologies like the Office Graph may improve further still.

Many 'on-premises' organizations moving to the cloud keep the mailboxes of leavers. The final cornerstone of the compliance story for keeping all mail data within the Microsoft service is a feature called Inactive Mailboxes. Rather than continue to pay for an E3 or Exchange Online Plan 2 licence for people who have left the organization, Inactive mailboxes allow the mailbox data to be kept for as long as the hold policy requires for free. The user must be deleted from Office 365 and will not be able to access the data, but all data is still discoverable and can be re-attached at a later date if required.

Today

These and ongoing improvements, such as ensuring hidden headers, e.g. BCC'd recipients, are captured, have meant that over the course of the last 12 to 18 months, records-keeping worries are unlikely to be a real blocker for a move to Office 365. Across almost every industry vertical Microsoft has case studies providing evidence that the built-in functionality works, and via the *Trust Center* links are available to information on compliance with global regulations.

The Challenge for Organizations Moving to the Cloud

Where a decision is made to move entirely across to Office 365, a successful project needs to clean up what is left behind.

Continuing to run a legacy archive system is costly both in hardware and software maintenance costs – especially if it continues to be used and requires updating.

Even leaving PST files behind on user workstations should be classed as only doing half a job and will eventually ‘bite’ an organization at a later date, either due to incompatibility with the client at some point or the more likely – data loss.

Likewise, if a legacy archive platform is in place then choosing a migration product for moving and organizing the data should not be an afterthought.

Some organizations who would prefer to manually export and import archives find it actually extends the project much further than expected, as well as find it harder to access budget for the right tools for the job after the initial migration is complete.

In the same breath, organizations looking to make best use of Office 365 are looking for a return on investment and the appetite for a large sub-project for legacy archive or PST migration is not there for many mid-size to smaller enterprise customers. A quick, relatively pain-free migration is the order of the day.

Moving the data back into Exchange, whether on-premises or in Office 365, is quite complex even without taking into account that some items may be very difficult to retrieve and that the workflow for moving the data needs to be flexible.

Organizations need to make decisions about how to deal with data for problems such as:

- What to do when moving items long-deleted by a user back into their mailbox,
- Workflow for moving data from journal repositories into mailboxes, or
- Where to move messages relating to expired shortcuts to; or
- What to do with mailbox content for people who have left the organization.

The answers are often for each organization to make but usually are aimed at maximizing the use of Exchange online archiving, In-Place Hold and inactive mailbox functionality to help ensure users are not inconvenienced. The aim should be that the business finds the new solution an improvement over the previous third-party solution and the organization gets best value from Office 365.

Migration workflows also need to take into account challenges that might not be present on-premises, like bandwidth availability. It is not uncommon to need to migrate terabytes of data into Office 365. The process for moving data needs to take into account these limitations and afford the opportunity to use disk-shipping technologies, either between sites or directly into Microsoft datacenters.

Finally, and above all else – the migration process must ensure the integrity of the data moved. Rigorous compliance demands across the globe often mean that when data is migrated between systems there may be a point that an organisation is expected, when supplying data as evidence, to prove that when the archive was migrated, the correct chain of custody occurred. Tests to ensure that the source and target items remain the same must occur and equivalent policies are applied to items.

Steve Goodman is a consultant helping customers deploy and adopt Exchange, Office 365 and Lync. He is actively involved with the Exchange community, authoring, blogging, speaking at conferences and hosting a bi-weekly podcast, The UC Architects. Steve holds multiple certifications and is an MVP (Most Valuable Professional) in Exchange Server. Steve blogs at his personal website www.stevieg.org



FACT Electronic evidence can be used in court to convict persons of crimes.

FACT Electronic evidence is easily altered, deleted or just plain 'lost'.

These facts make it vital for an organization to have tight and meticulous control over the handling of electronic evidence as it is collected and processed in response to an eDiscovery request. *Without the relevant assurances there could be allegations of tampering or misconduct (which could compromise the outcome of a case).*

Now let's go back a step or two...

Before an eDiscovery request even surfaces, you should have a consistent governance policy in place over your electronic records and their life-cycle.

More specifically, you must have a policy in place that manages and articulates the retention and disposition of data in accordance with your business and legislative remit.

Additionally, protecting the integrity and eDiscover-ability of your data and records throughout their entire lifespan, could mean storing and maintaining the accessibility of your data for **several decades**.

During the life of an email record, change is always inevitable:

- The device on which it is stored may become obsolete or deteriorate.
- The software used to manage the record may become outdated and no longer maintained by the vendor.
- The format of the email or attachment may become difficult to support.

Organizations have a duty of care - and a legal remit - to ensure they have created the best processes for keeping data secure yet readily available and discoverable. *Not addressed properly, this area could hold organizations and individuals accountable in a court of law.*

If a physical move of electronic records – whether to a new storage platform or 'the Cloud' - has compromised their integrity, reliability or completeness, any evidence produced in response to a future case could be overturned and may lead to suspicion of deliberate spoliation, enormous costs and increased levels of scrutiny.

E.g. The US Department of Labor's Occupational Administration requires that some health-related records be kept for either 30 years or the duration of a person's employment plus 30 years.

As demonstrated by the high profile and very public 'phone hacking case at News International, the excuse of a botched shipment of archived emails to India was viewed dimly as an excuse for losing emails relevant to the case, and the subsequent discovery of an intact archive led to 'no stone being unturned'.

So what's important as you move your data to the Cloud?

10

THINGS TO CONSIDER WHEN PLANNING TO MIGRATE LEGACY ARCHIVES TO MICROSOFT OFFICE 365

Never underestimate what's involved in a migration. Although some migration vendors may paint a simplistic picture that focuses on shifting large data volumes at high speeds, it's worth taking a step back and looking at all aspects of a migration.

We've seen several migration projects come to a screeching halt when the needs of the legal department, end users and the business as a whole have not been taken into consideration.

There are also fundamental differences in the way Exchange and email archives store data versus Office 365 which, if not properly addressed, can later invalidate a compliance-led migration. *In other words, 'bite you in the rear'.*

1

WHAT SHOULD YOU TAKE?

If you've been maintaining a journal archive to date, it's pretty much a 'no-brainer' you'll want to move it all – perhaps with the exception of a date-based cut-off point that coincides with your corporate retention policy.

If you haven't been journaling, yet have been archiving mailboxes with a view to meeting compliance and/or business needs then, again, you should be migrating in line with your retention policy.

If you've purely been archiving as a practical step to take the storage strain off your Exchange servers, don't assume that practicalities alone should determine what is moved.

It is not uncommon for the IT department to attempt to limit the amount of data they migrate in order to shrink migration times and costs. However this is a flawed strategy for at least 2 reasons:

1. In the event of a future litigation, mass deletion of emails to simplify a switch to a new email system is not viewed as a good excuse for losing data. At worst it could be viewed as deliberate spoliation.
2. Overlooking the needs of end users - who have relied on readily being able to search and retrieve from their archives in order to carry out their job - could have a serious impact on your business' bottom line. It will also drive the perception that the migration was not well-executed or a benefit to the business.

If you haven't to date enacted a formal information retention policy for your business that relates to email (and other content), the point of migrating to Office 365 is the perfect opportunity to get one in place.

It is also the perfect opportunity to **share the responsibility** of policy management outside of the IT team, thanks to the enhanced management console provided by the new Office 365 Compliance Center.

Your retention (migration) policy should ideally be simple, clear-cut and based on headline data such as sent date and custodian. *For example, 'Keep all email belonging to staff in the finance department for 7 years, and everyone else's for 2.'*

Of course, a broad-brush policy like this could mean you end up migrating a lot of 'rubbish', so how do you avoid this?

One approach is to consider looking at user-applied classifications or advanced filtering techniques to drill into the meaning of email content in an attempt to identify the emails that are of use to the business. Such techniques *can* lead to unsatisfactory results. For example:

- Users may not have correctly filed or tagged all relevant items. *Bear in mind typical users are not 'traditional records managers' - we find that only a small percentage of staff receive clear guidance on corporate records management needs.*
- You could inadvertently exclude seemingly 'valueless' emails that may turn out to be evidence in a future litigation case. *Who's to say that repeat invitations to lunch might not be relevant in a sexual harassment case?*
- It could take your legal department **weeks** to agree on what constitutes 'relevant content' (thus eroding any time-benefits gained by reducing data volumes).

Ultimately your policy on what to migrate should be informed, clear-cut, defensible and fully in line with your legislative requirements, risk profile and your business needs.

Whatever approach you use to limit what you take, it shouldn't be just a one-off 'moving house clear out' that won't be carried out consistently in the future as an ongoing process.

2 | WHERE SHOULD YOU PUT IT?

So now you've decided what you want to take, the next question is, 'Where in Office 365 are you going to put it'?

If you're moving Journal archives, these need to be put into a very specific place. Read more [later in this document](#).

If you're moving users' archives into Office 365, it's important to take a **people-centric approach**, even where compliance is a primary goal. Here are some things to consider:

Ensuring a great user experience as you migrate is vital.

Bear in mind that third-party archives typically work by replacing the original email with a much smaller shortcut (stub) that 'looks and feels' like a regular email. In fact users are often unaware they are working with an archived email.

If emails become difficult to locate, or disappear post-migration, this can have a serious impact on individual productivity and the bottom line of the business as a whole.

Careful consideration around where in Office 365 you migrate emails, and clear communication to users on where they can find their emails post-migration is therefore critical to success.

Suggestion: Move archived data into Office 365 Online Archives. If your legacy archive uses a simple age-based policy such as 'archive everything > than 1 year', and you plan to continue this using Office 365 Messaging Retention Management (MRM) policies, you could move users' archives directly into their online archive (more commonly known as the In-Place Archive) and advise users accordingly.

It's not always this clear-cut. Your archive policies might not be suitable for this approach. For example:

- If you're archiving user mailboxes to meet compliance needs, you may be archiving everything on a daily basis.
- If you're archiving mailboxes to minimize Exchange sizes, you might automatically archive emails over a certain size.

In each case users will expect to see these 'younger archived items' in their Office 365 primary mailbox 'post migration'.

BEWARE: putting 'prematurely archived' items directly into the In-Place Archive, could result in 'end user confusion'.



Will emails end up in the right folders?

As you migrate the contents of users' archives, the same folder structures should be maintained.

This can be challenging given that some archive systems don't consistently track when users delete their shortcuts, re-folder them, forward or share them with co-workers. If you do not pay close attention to this level of detail in your migration you could find yourself with:

- Emails that reappear in Office 365 when the user has deleted their shortcut.
- Emails that unknowingly get migrated to the wrong folder within Office 365.
- Emails that fail to get migrated to all relevant people within the business.

Scenarios such as this will confuse end-users, as well as impact productivity and put a load on the help desk.

Additionally, having emails 'end up' in the wrong place post-migration can have adverse information governance consequences, so check that these issues will be addressed as you migrate.

Also remember your migration strategy for user archives will need to include the removal of legacy archive shortcuts, either pre or post-migration.

As you migrate to Office 365 these shortcuts will effectively get replaced or 're-hydrated' with the original email.

2 | WHERE SHOULD YOU PUT IT?

Continued...

Deleted Items

There may be archives that aren't to be accessible by end users, but you still need them to be preserved and searchable to authorized personnel. *For example, items that have been deleted by users yet remain in the archive.*

To preserve these emails in the correct way you'll need to migrate them into a special hidden subfolder within the [Recoverable Items folder](#) (RIF). *Note that there's a RIF associated with the primary mailbox and the In-Place archive, and you should check your chosen migration method supports migration into either as needed.*

In all cases the mailboxes you are migrating into in Office 365 should be put on permanent In-Place Hold in order to maintain the same protection against deletion that you had with your third-party archive environment.

See also point 4 on protecting records on hold.

'Leavers' archives

If you have a need to preserve archives belonging to staff no longer with the company you'll have to provision a mailbox in Office 365 for **each** individual and migrate their data into this.

But don't worry – this won't be as costly as you think.

Thanks to Microsoft's [inactive mailboxes](#) facility, it's possible to commission leavers' mailboxes, migrate archives into them, put them on In-Place hold, delete the mailbox and then re-assign the associated mailbox licences.

Meanwhile the contents of leaver's archives will remain available indefinitely for eDiscovery.

Note that Microsoft expressly prohibits the use of a single user's In-Place Archive for storing items belonging to multiple users. This rules out the concept of grouping leavers' archives into 'departmental mailboxes'.

Practicalities

Although the Enterprise Office 365 plans now offer a generous 50GB primary mailbox (which could easily accommodate most users' archives), there's practical reasons behind migrating older items into the In-Place Archive.

For example, if your organization is planning to take a 'hybrid' approach to its overall Office 365 migration, users can continue to work with their primary mailboxes 'on-premises', meanwhile their archives can be migrated into cloud-based In-Place Archives 'behind the scenes'. With this approach, there's no interruption to archive access.

Also, if you plan to use 'full Outlook' (desktop) clients with the ability to access emails when working offline, having very large online mailboxes to synchronize onto local systems will create a huge network overhead during the migration process. This will require that you continue to do daily synchronizations which will take a while.

If you use the latest Outlook client there's a slider to control how much gets synchronized, but older versions of Outlook may struggle to provide a performant synchronization service.

In summary, simply migrating everything in your archive into the In-Place Archive is not necessarily going to be the best approach. Making sure you understand what archive policies you have in place already, why you have them in place, and what your future policies should dictate will ensure that the results of the migration will align to all the needs of the business, inclusive of technical, end-users and your compliance-focused users.

3 PRESERVING CHAIN-OF-CUSTODY

Any time an electronic record is moved between storage devices or locations, there's a potential point-of-failure, weak link, or a possible disruption that introduces a number of risks, including deletion, alteration or substitution.

It is this risk that can compromise the integrity and reliability of your data - or what's considered as material evidence - called upon in the event of litigation.

Technically speaking, chain-of-custody can be strengthened by minimizing the number of conversations, custodians, and interim stages which the item(s) being transferred go through. Ideally, you should look for a migration process that moves your data directly from your on-premise repository to Office 365 within in one audited, end-to-end transaction.

If the available bandwidth to the Cloud, combined with the amount of data to be transferred, means this direct approach is not possible, an Azure-based transfer or Drive Shipping using interim files may need to be used, but be aware that this multi-step migration may compromise chain-of-custody without the appropriate security and activity alerting mechanisms in place.

Regardless of the approach taken, it is vital to be able to prove that:

- All relevant evidence has remained unchanged.
- All data has been successfully tracked through to receipt by the next stage or custodian (for example, with detailed transaction records including a timestamp and the ID of the item in both the source and the destination).
- All migrated data has been fully accounted for at all stages.

Having these assurances means that *potential* evidence has been handled in a manner which allows no doubt that it could have been accidentally or deliberately altered or substituted. ***It also gives peace-of-mind to business stakeholders as they hand over mission-critical data to a third-party service.***



Knowing the current location of evidence is not enough; there should be accurate logs tracking the movement and possession of evidence material at all times during its lifecycle.



R. Yeager 'Criminal Computer Forensics Management', InfoSecCD, ACM, Kennesaw, USA, 2006

4 PROTECTING RECORDS ON-HOLD

If one or more litigation cases is running during the time of your planned migration it is imperative that any ring-fenced data sets (i.e. data that has been put on legal hold to prevent deletion) is preserved and that legal hold is maintained once transferred to the new system.

Likewise, if you have any existing retention policies, you may wish to map these into the equivalent retention tags that will override default retention policies when you migrate this data into Office 365.

5

HANDLING FAILURES & EXCEPTIONS

A lot can happen over the life-span of an archive: indexes and storage devices can get corrupted; not thoroughly-tested or bug-ridden upgrades may be applied and 'foreign' data could find its way to your archive from outside source or systems.

As a result, it's not unusual for emails to 'get broken' and fail to migrate.

Your planned migration route should, as a minimum, perform a healthy amount of baseline checks to test that all items can be reliably searched and opened post-migration, and that you are not just moving 'broken' emails and/or attachments that may be unusable in Office 365.

Ideally there should be a mechanism for automatically retrying any failures (which may occur owing to temporary environment issues), and in the event of a 'permanent failure', a detailed report, inclusive of explanations to assist with trouble-shooting, should be provided.

As part of determining your optimum migration strategy we recommend that you get input from your legal department asking them to consider their desired course of action in the event of a failure.

An example of this would be asking if it is acceptable just to log a failure, or does each failed item need to be investigated fully?

Even if it seems that an email is damaged beyond repair, it's important to note that part or all of it may still be salvageable. In the situation when presenting evidence in a court of law, it's extremely important to be able to prove or illustrate that the business took all of the necessary precautions and steps during its migration to be able to reconstruct these objects.

Keeping in mind that remediation of failed items may demand additional time and funding, being prepared will allow you to investigate the options that exist to ensure you are prepared for a winning scenario.

6

MOVING LEAVERS' DATA

The new storage model in Office 365 requires that you provision a mailbox for all users, *including staff that have left the organization.*

Where co-workers or departmental managers need to maintain immediate access to former employees' email records in order to do their work, it is best practice to migrate these legacy archives into [individually licenced mailboxes](#) and then set appropriate delegate access permissions.

If this data is required purely for eDiscovery by authorized individuals such as compliance officers, or records managers, you can achieve this without the expense of dedicating a *permanent* mailbox licence for each departed user.

This is achieved by migrating the data, putting the corresponding mailboxes on 'In Place Hold' and then making these mailboxes [inactive](#).

If you plan to move journal archives into Office 365, this will invariably include leavers' data that needs handling similarly. [Read more later in this document.](#)



7 | MANAGE WHAT YOU LEAVE BEHIND

If your aim is to make Office 365 a one-stop shop for Information Governance and eDiscovery, you should **avoid having ancillary data repositories as these will need to be managed and included in any future eDiscovery case.**

This means, following successful migration of your legacy archives to Office 365, it is best practice to comprehensively manage the disposition of emails 'left behind'.

Organizations that believe they have 'defensibly deleted' records *excluded* from a migration may be surprised to learn that other viable (and therefore discoverable) copies of their archives are likely to exist elsewhere.

There is no such thing as a naïve operational court today - they are all 'in the know' when it comes to discoverable data.

For example, archive backups may not have been properly rotated, and there may well be a short time-window where deleted data is still be recoverable, or there may be an off-site copy of an archive that has been overlooked.

PST files are a monster issue to contend with: They are notoriously difficult to track, especially difficult to search at a content level, and there's the possibility that they may exist on someone's external drive or USB. It's almost impossible to ensure that your centralized deletion and migration policies marry up with all known users and their own data archives.

See the next section for PST migration best practices.

Any strategy that involves migrating just the data 'deemed necessary', and leaving the rest behind to 'age in place' in PST files or legacy archives is counter-productive, as these repositories will be considered 'fair game' in the event of a litigation case. As a consequence, you'll face the costs of discovering against these repositories as well.

8 | MINIMIZE MIGRATION TIMES

Depending on how long you've been archiving for, your email retention periods, trends in staff 'churn', and archive storage compression ratio, the volume of email to be migrated from your archives will be orders of magnitude larger than that in your 'live' email stores.

This makes raw speed a critical factor in your migration strategy – even more so if you operate in a highly litigious industry sector where you need to be prepared to respond to a request in a short time-frame.

If you're migrating user archives, you'll also be keen to minimize the amount of time that users are separated from their data, as this could have a severe impact on their productivity and your help desk.

HERE'S SOME POINTS TO CONSIDER:

- Your chosen migration route should feature cutting-edge performance techniques, both with respect to ingestion into Office 365 and extraction from your source archives - but never at the expense of the integrity or completeness of your data. Keep in mind, that although direct archive extractions can be 10x faster than using vendor supplied API, direct extractions may be unable to access emails on tiered storage options that are reliant on the API for access.
- Direct end-to-end migrations will give you optimal chain-of-custody by cutting out interim steps. However, where network bandwidth is preventing you from reaching the speeds you need, you have the option of PST-based uploads via Azure, or Drive Shipping, *but look for maximum security assurances and auditing as both these approaches risk breaking chain-of-custody.*
- As outlined in point 1, if you use filtering to reduce the amount of data you move, make sure your filtering matches your retention policies and is as simple and defensible as possible. Bear in mind that having your legal department spend time drilling into the 'meaning' of data and determining whether it is relevant to your business has the potential to massively 'put the brakes' on your migration. It can also compromise the validity and completeness of your migrated data set.

9

DON'T NEGLECT PSTS

Personal storage tables - more widely known as PSTs - have long been a support, storage, information governance and eDiscovery nightmare for enterprises.

A switch to Office 365 - with its large storage capacity - is the perfect opportunity to eliminate this outdated way of providing a mailbox 'overspill' facility.

What's more, by migrating PSTs to Office 365, organizations can get to perform eDiscovery across their contents – something that is practically impossible to achieve with PSTs located on individual hard drives.

There's some huge challenges in migrating PSTs, however, especially in organizations that have many 1,000's to tackle.

COMMON HURDLES CAN INCLUDE:

- **Locating the 'Crittlers':** The task of hunting down PSTs, migrating them to Office 365 and being able to track their migration, can be a highly complex undertaking that requires significant 'horsepower' and sophisticated functionality.
- **Filtering What You Take:** The nature of PSTs is that one physical file contains many individual emails. For this reason, the Microsoft PST Capture tool and third-party tools typically migrate the whole PST, which could mean importing lots of emails that fall outside of your retention policy.
- **Eliminating Duplicates:** PSTs are easily copied – both by over-zealous users as well as PST backup routines. As a result it's easy for migrations to end up with many duplicate emails.

Management and migration of PSTs will be the subject of a separate white paper in this series.

In short, these and many other challenges dictate that PSTs are often excluded from a migration project.

A common approach for an organization to take when attending to a compliance remit is to prevent users from writing to PSTs or creating new ones, and allowing them to 'age in place', with the understanding that:

- PSTs remain difficult to search in the event of an eDiscovery request.
- PSTs can only be managed according to the overall age of the PST file, not the individual items contained therein.

The ideal scenario when dealing with these problem files is to be able to migrate their content to Office 365 in accordance with your retention policy and at the individual email level, with the ability to filter out duplicates.

Compliance issues aside, by migrating PST content into Office 365, users are able to access their data from anywhere and from any device – without the restrictions and risks associated with locally stored PSTs. This represents huge productivity and mobility benefits.

Microsoft states that the use of Personal archives (PST files) is unsupported over a local area network (LAN) or wide area network (WAN) link. Additionally they say that 'PST files are not meant to be a long-term, continuous-use method of storing messages in an enterprise environment.' [Read more.](#)



10

TREAT JOURNAL ARCHIVES WITH CARE!

Email journals aren't just about capturing a copy of every email sent and received. They also preserve valuable **metadata**.

Metadata – data that describes data - is deemed to be highly relevant in the process of eDiscovery. Not only can it provide valuable contextual information, it can speed up the initial **collection** stage, as many search engines create a light-weight index of available metadata, enabling a 'quick first pass' that does not need to drill into content.

The Exchange journal service (by default) preserves otherwise hidden metadata relating to **ALL the recipients of an email**.

This includes any recipients that were blind copied (BCC'd) or any recipients that were part of a distribution list (DL) at the time the email was sent. *Again, keep in mind that the members of a DL will change over time as their roles change and as staff leave or join the organization.*

Any migration that fails to preserve the metadata captured by journal archives would mean that a search for all the people that were 'party' to financial misconduct or other nasty business conducted over email would be incomplete.

Up until now, if you wanted to retain your existing journal archives and keep running a journal archive going forwards, you would need to maintain your existing archive/journal service on-premises or look towards using a third-party journal service.

Either way, you'd have two locations to maintain and search in order to meet information governance and eDiscovery needs.

The good news is that now, although very different to the Exchange journal model, there are new facilities that have been added into Office 365 which are designed to replace the role and functionality of the Exchange journal service.

So if you plan to migrate an Exchange journal, it's worth understanding what needs to happen 'behind-the-scenes' of your migration so you can rest assured that your journal migration approach is the proper one.

Migrating your journal archive requires more than migration know how, or hopeful thinking – it's a subject that requires further exploration therefore we dive a bit deeper into this topic in the next section of this whitepaper.

Replacing Journals With Office 365

The Exchange Journal mailbox was designed to enable organizations to capture a copy of all or selected email traffic in accordance with regulations including Sarbanes-Oxley, SEC 17A-4, HIPAA, the Data Protection Act.

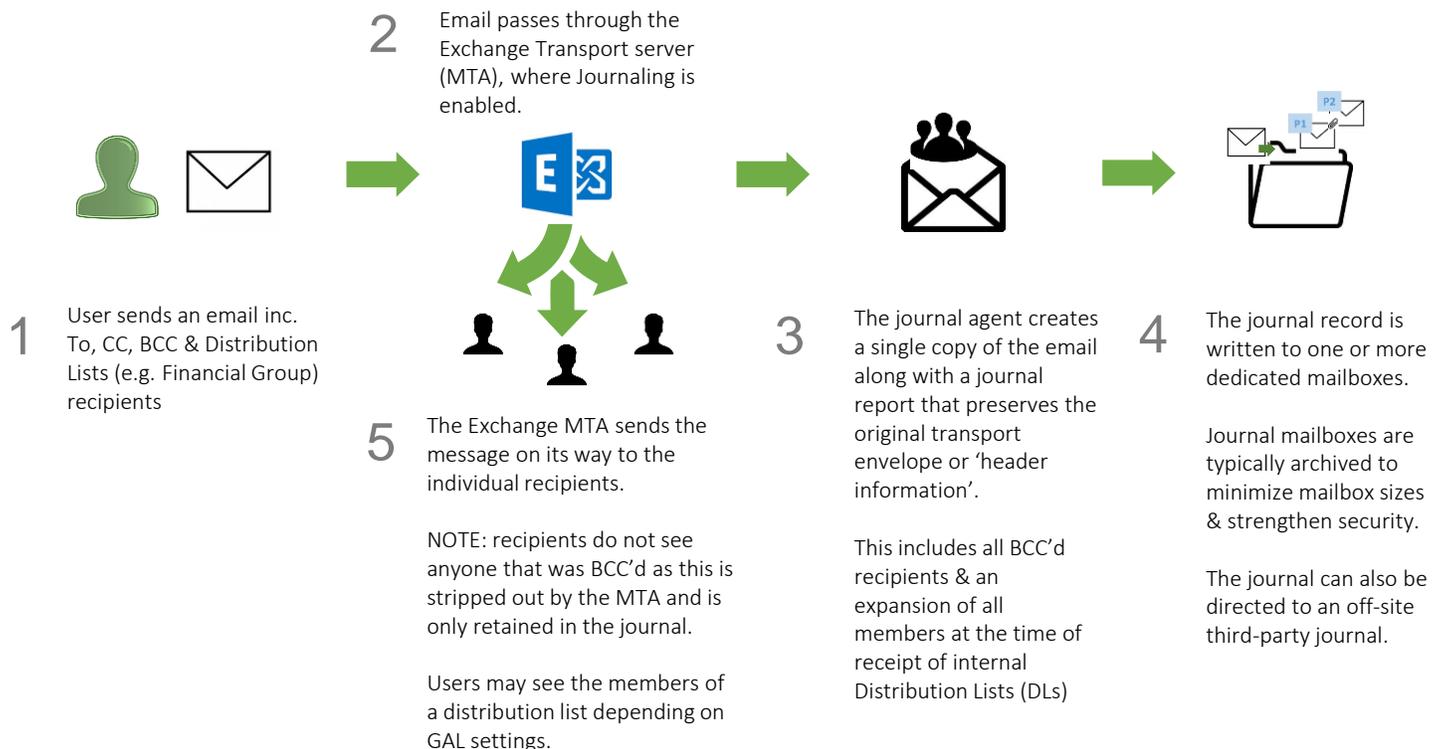
In brief, when handling inbound or local traffic, the Exchange Message Transfer Agent (MTA) - depending on the rules you configure - directs a **single** copy of each email it handles to one or more journal mailboxes, along with a copy of the 'message envelope' it uses for delivery purposes.

The **envelope** provides a record of everyone that would have received a copy of the email, including any recipients that were part of a local distribution list (DL) at that point in time, and anyone BCC'd on the email. *The latter is automatically stripped off by the MTA prior to delivery so as not to be visible to the individual recipients.*

From a compliance perspective, if this full set of recipient metadata is not captured, the initial data collection step of any future eDiscovery exercise could miss out some of the relevant people that were intended to receive an email. Therefore anyone BCC'd on an email could potentially avoid detection. Likewise, staff members that *used* to be in the Financial Trading Department, could be excluded from an investigation involving past insider dealing if they no longer appear in the current DL.

To add to the overall complexity of preserving this information when migrating, journal mailboxes are often put into secure third-party archives, and each vendor had its own *proprietary* way of preserving envelope data alongside the corresponding email.

The Exchange Journal Model



How Office 365 Replaces Journal Functionality

To re-cap, Microsoft Office 365 has enhanced its Office 365 model to achieve the **same ‘compliance outcome’ of a journal service**. Here are the main features to understand:

- Instead of using a large, centralized, single-instanced mailbox that is inherently difficult to scale and failover, Microsoft has leveraged its optimized multi-instance storage model. This allows each user to retain his/her copy (journal) of all emails sent/received with zero performance penalty and no single point of failure.
- By putting all relevant mailboxes on Litigation or In-Place Hold, all emails sent and received are retained indefinitely.
- Even if a user deletes an email, the email gets removed from the user’s view, but is moved into a special hidden folder inside the Recoverable Items Folder (RIF), where they are available to the eDiscovery process.
- Any BCC’d recipients will be retained indefinitely in the senders’ mailboxes.
- The members of any distribution lists (DLs) are expanded at the point of sending and stored in hidden headers in senders’ emails so they are fully discoverable.
- Inactive mailboxes (i.e. those belonging to leavers) can be put on Indefinite Hold without a license penalty.

So How do You ‘Map’ Legacy Journals into the Office 365 Model?

In order to migrate all the legacy data captured by the ‘old’ journal format to the new Office 365 model, several things need to be addressed.

These include:

- **Multi-instancing:** By this we mean that the single-instanced journal store needs to be converted back into a multi-instanced data stream comprising a copy of the original email for each recipient listed in the original email envelope. This sounds like a huge overhead, but done the right way, the volumes involved and the impact on mailbox quotas need not be an issue. Also remember that Office 365 is optimized for this multi-instanced storage model.
- **Handling Leavers:** Your journal archive will naturally involve emails exchanged by staff that are no longer with the organization. Although the new Office 365 model requires you to provision a mailbox for each leaver, by using the Inactive Mailboxes facility you can migrate leavers’ legacy journal emails, close off their mailbox, and if handled in the proper way, you can re-use the mailbox licences for your ‘live’ users.
- **Handling Deleted Items:** When migrating journaled emails belonging to staff still with the company, it’s possible you’ll be migrating emails that have long since been deleted. For this reason you need to move them into a hidden area. I.e. the designated folder in the Recoverable Items Folder.

- **Preserving BCC’d data:** It’s worth remembering that only the sender of a message sees any BCC’d recipients. As a result, you need to make sure that two different versions of each message are pushed into Office 365 - one to the sender, which includes the BCC data (and is therefore fully discoverable) and one to the other recipients which does not include the BCC data. Without this, the confidentiality of BCC recipients is broken during migration.
- **Preserving Distribution Lists:** Any historic distribution list (DL) information needs to be mapped into the new Office 365 hidden header field for the sender’s version of the message.

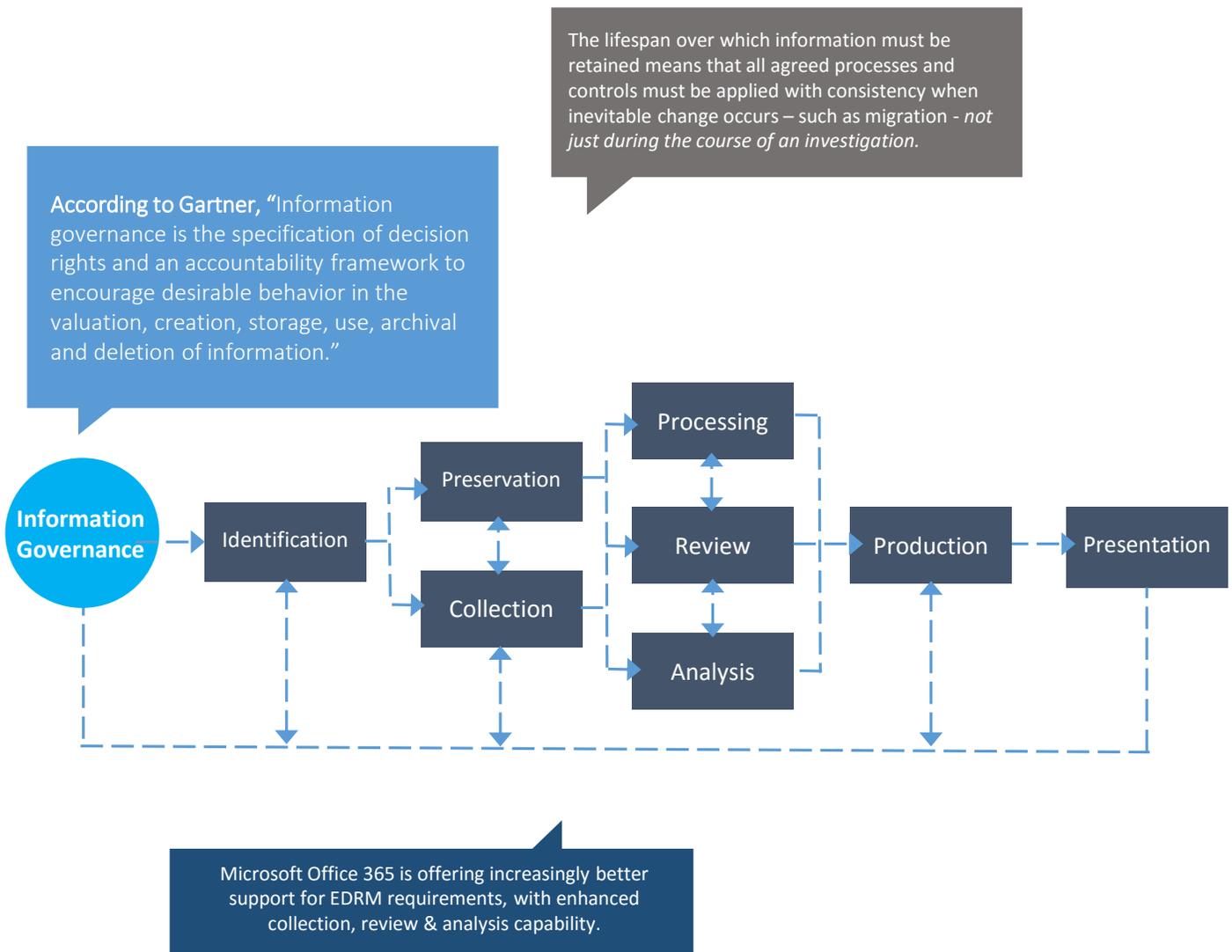
By addressing these areas correctly, you can be assured all the relevant data is not only protected, but that it is IN THE RIGHT PLACE as far as the eDiscovery process is concerned.

In addition to correct mapping of data, there’s two overarching factors to consider when moving journal archives owing to their sensitive nature and their size, namely:

- Journals must be moved with care
- Journals must be moved with speed.

For more information on each of these areas refer to our earlier sections on chain-of-custody and performance.

The Electronic Discovery Reference Model - EDRM



Enhanced Office 365 eDiscovery Features

Other significant developments to the Office 365 compliance and eDiscovery offering include:

Putting Responsibility into the Right Hands

In the past, the information management and eDiscovery features available in Office 365 would need to be managed by IT staff owing to the technical nature of interfaces such as the Exchange Administration Center (EAC).

However, this is rapidly changing with the introduction of the new easy-to-use Office 365 Compliance Center.

The **Compliance Center** is designed to take the onus off the IT department with a new, friendlier interface designed for use by key specialist roles—such as compliance officers or HR personnel.

At the time of writing it includes the ability to:

- Define retention management policies determining rules for archiving and deleting users' data.
- Manage the types of mobile devices that can access the Office 365 service including the ability to block or wipe devices.
- Review service configuration audits that include all administration activity, any 'holds' placed on data, any changes in administrative rights, etc.
- Perform eDiscovery across content from Exchange Online, SharePoint Online, Lync Online, and even OneDrive shares.

Other activities currently managed in the EAC, such as data leakage prevention policies and applying In-Place and Litigation Holds, will be moved across into the Compliance Center over time.

Ideally the process of eDiscovery and other compliance activities should be handed over to the HR or legal/auditing department, and should not fall within the remit of the IT department. Not only will this remove an administrative overhead, it will avoid IT staff from being involved in potentially sensitive situations.

Enhanced Searching

'Background' enhancements to the eDiscovery service are constantly being made. This includes the ability to now conduct 2 concurrent searches, each across 10,000 mailboxes, as well as carry out federated search across an unlimited number of SharePoint Online and OneDrive for Business sites—in a single eDiscovery query.

Advanced Analytics

The remarkable increase in electronic records stored by enterprises is creating a huge burden as they struggle to provide a cost effective approach to the first stage in any eDiscovery process: *collecting, processing and preparing relevant data that may be required by litigation.*

The [recent acquisition of Equivio](#) and its market-leading Zoom product, means Microsoft will address this need.

Viewed in the context of the EDRM model, where the current Office 365 compliance capability addresses the information governance, data preservation and collection stages of the eDiscovery process, the new service will **massively streamline** the analysis of large amounts of unstructured data extracted from Office 365 and other sources, enabling organizations to focus in on **just the data that is relevant to a case.**

As well as reducing data review and preparation costs, the service will help organizations with what is termed '**Early Case Assessment**', by enabling them to gather enough relevant information to quickly determine whether it is worthwhile to either prosecute or defend, gauge how much it might cost to fight, or indeed, make a decision to settle a case out of court.

Hosted as a web-based Azure-based service the new Microsoft eDiscovery solution will include a range of **advanced techniques** designed to save time and subsequent review costs when handling and culling large data sets. These include:

- De-duplication and grouping of near de-duplicates.
- Reconstruction of email threads and removal of redundant repeated content.
- Semantic analytics to group together items that center around high level 'themes'.
- Key word search.
- Predictive coding and statistical analysis techniques that involve initial training using sample documents provided by your legal team.
- Collection of related documents that feature similar concepts and terms.

These services significantly reduce the time and costs faced by organizations in preparing for litigation or dispute resolution.

Summary

- The new information governance and eDiscovery capabilities offered by Office 365 – combined with the potential to shift the onus of policy implementation from technical staff into the hands of legal, records and HR personnel – offers many compliance and financial benefits.
- The costs of migrating legacy data into one place – i.e. Office 365 - should be weighed up alongside the overheads and complexity of maintaining and searching email records held in multiple distributed data stores.
- Identifying what data should be moved, and understanding **how to map this data** into the new Office 365 model so that it is fully manageable and discoverable ‘post migration’ should be addressed fully in advance.
- The migration process should take into consideration the information governance needs of the company, with close attention paid to moving the email, the attendant metadata, retention policies and other vital attributes.
- Any strategies for minimizing what you migrate should be in line with existing information management policies. If you don’t already have an information management policy in place, this is the ideal time to get your house in order. Your policies should be applied consistently from here on in and not just ‘one-off special’ policies for your migration.
- The process of setting retention policies should include all relevant stakeholders, ideally IT, legal, business and compliance staff that have information management experience.
- Generally speaking, simpler, more clearly-defined policies that rely on ‘headline metadata’ tend to be easier and quicker to implement and easier to defend versus using solutions that automatically make retention decisions based on the meaning of email content or that rely on end users.
- Proving that your migration has been carried out with all due diligence in maintaining the integrity, completeness and viability of the data, so that any future eDiscovery exercise is not compromised, is vital.
- Although storage is practically unlimited in the Cloud, retaining excess data makes the process of eDiscovery more time-consuming and costly, and may risk providing litigants with more information than they’re entitled to receive. With this in mind, make sure you have an effective policy for managing the deletion of data.
- Take advantage of the fact that Office 365 offers a platform for putting information retention management into the hands of the **right personnel** in your organization!

By following these recommendations you’ll be able to reduce the ongoing costs of managing and discovering against legacy data by putting it all in one place: Microsoft Office 365.



Since 2007 TransVault has led the market with its highly specialized archive migration solutions for the enterprise.

When businesses encounter technological change brought on by a merger or acquisition, a planned shift to the Cloud, or the obsolescence of their archive or storage solution—they turn to TransVault to preserve accessibility to their business records.

Over 1,000 customers from around the world have relied on the TransVault architecture to protect the integrity of their valuable legacy data whilst ensuring chain-of-custody, faultless eDiscovery and seamless user accessibility—no matter the complexity of the migration, nor the archive platform.

TransVault solutions are available through an accredited international network of partners with dedicated Migration Specialists certified on TransVault solutions.

TransVault continues to achieve year-on-year growth and has become the preferred archive migration solution for global customers in all verticals, especially those with a heavy dependency on data sanctity and regulatory practices.

Supported Platforms:

- Autonomy Consolidated Archive (ACA)
- Autonomy Message Manager (formerly CA Message Manager)
- Autonomy NearPoint (formerly Iron Mountain/Mimosa NearPoint)
- Autonomy ZANTAZ Digital Safe
- Autonomy ZANTAZ EAS for Exchange & Notes
- EMC EmailXtender for Exchange & Notes
- EMC SourceOne for Exchange & Notes
- Google Apps Vault
- HP RISS.IAP for Exchange & SMTP
- IXOS-eCONserver (Open Text)
- Ilumin Assentor
- Metalogix Archive Manager (formerly Exchange@PAM)
- Microsoft Exchange on-premises archive
- Microsoft Office 365 & Exchange Online Archives
- Dell (formerly Quest) Archive Manager
- Veritas (formerly Symantec) Enterprise Vault for Exchange & Notes
- Unify/Daegis Central Archive (formerly AXS-One) for Notes
- Cloud-based journal services including Mimecast, Enterprise Vault.cloud & Proofpoint
- On-premises journal services/appliances
- PST, EML, MSG & NSF files

For more information on TransVault products or your nearest TransVault partner:

North America: 646 808 0407

Europe: +44 (0) 3333 404433

General Enquiries: info@TransVault.com

Product Information & Licensing: products@TransVault.com

www.TransVault.com