EXONAR

The ultimate guide to data migration

All the tools and tactics to help you migrate the right data safely and securely, and how to do it

anacomp®

96% of companies have migrated to the cloud[1]

Whether it's Microsoft 365 and Azure, AWS or others, the trend for cloud migration is set to increase in 2021. As the pandemic took hold, people were directed to work from home wherever possible, which accelerated digital transformation exponentially to drive the mass cloud migration witnessed in 2020.

But it was only ever a matter of time.

Today, every company is a data company. Every business, regardless of its sector, is producing extraordinary amounts of data that grows in size every day. We all operate within the information age, where data has become the most valuable commodity because there's so much value locked within it.

To take advantage of data, organizations are looking to adopt more digital tools, which involve migrating data lakes to the cloud. But the sheer size and scale of those repositories poses some serious questions:

▸ **What data is being migrated to the cloud?**

▸ **Why is it being moved?**

▸ **How will it be moved?**

▸ **What will happen to the data once it's in the cloud?**

▸ **How will data governance be maintained before, during and after the migration?**

Without understanding all the data the organisation contains in detail, any cloud migration will encounter problems, as there are always skeletons hidden in large legacy and unstructured data sets.

> ❯ **In over two-fifths of** migrations (41%) data issues caused the project to run over schedule and over budget[2]

In this report, we share insights from our strategic partner, Softwerx, a specialist in cloud security and governance. Together, we will show you how to gain visibility of your data at scale across the data estate to identify what you need to migrate and what you need to encrypt or delete.

> ❯ **About Softwerx**
>
> Softwerx combines its specialist knowledge and experience of Microsoft Office 365, Enterprise Mobility + Security and Azure infrastructure to help forward-thinking enterprises modernize and secure their businesses' data.
>
> **softwerx.com**

The biggest problem with data is over retention

Most organizations suffer from over-retention of data, with the main culprit being a fear of deletion.

We were recently joined by data security specialists, from leading UK companies in a roundtable discussion to identify their challenges. During the discussions, a senior information management specialist for a large fintech organization commented:

"Keeping [data] 'just in case' is a favourite phrase."

A lot of the time, organizations know they need to delete data, but are reluctant to do so:

"I found some data in Canada with 5 million people's information that hadn't been looked at for 10 years. Do we need it? Do people even know it's there?"

Over retention is often a source of immense frustration for data and security professionals. The European Data Privacy Officer for a bank specializing in wholesale assets said:

"When people ask if they can retain data, I can't give 'just in case' as a legal reason to do so!"
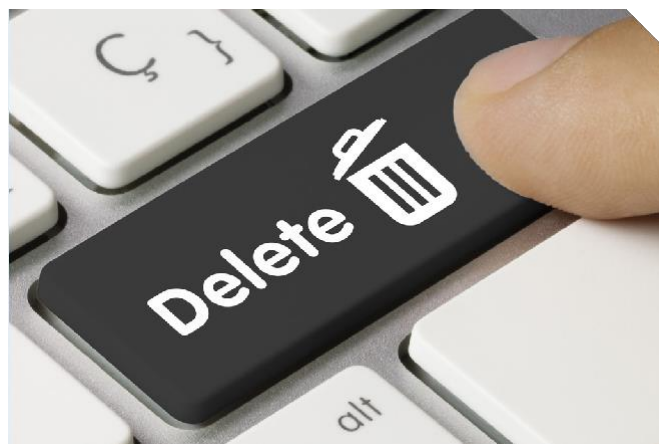
> Under Article 17 of the GDPR, organizations must legally delete data that is no longer required. It states, "the controller shall have the obligation to erase personal data without undue delay".

On a joint webinar with Softwerx, its technical data specialist, Adriaan Bekker, shared how every organization keeps data 'just in case'. But you have to decide to take action and do something about it. When his own organisation moved to Microsoft 365 and Azure, they took the decision not to migrate any data over 7 years old. Instead, older data was archived for one year. In that time, only one person asked for one document to be recovered.

> *"One of the biggest challenges we see in organizations is not having somebody to stand up and say they are prepared for their data to be removed because it's met the retention policy or the age at which it should be disposed of."*
>
> **Adriaan Bekker**
> **Softwerx**

You can't migrate it all

With data storage now relatively cheap, organizations might be tempted to just migrate it all – a simple lift and shift from the on-premise infrastructure up into the cloud.

But it's too risky, too expensive and plain unnecessary to migrate everything the business has to the cloud. Organizations that treat cloud migration as a simple 'lift and shift' of the IT infrastructure, over-provision by as much as 55% during the first 18 months and overspend by 70%.[3]

And then there's the risk factor.

All it takes is one mistake

Sensitive customer data saved in a spreadsheet on a local drive, a written password., a bunch of personally identifiable information (PII) left unencrypted - imagine taking that to the cloud – the risks would be huge.

Data security in a migration remains a top business priority for good reason – because a single small mistake can lead to lasting reputational damage that is costly to repair.

But everyone knows that there's data lying unprotected all over every business, leaving it exposed and vulnerable. That's why it's so important for an organization to ensure it has full visibility of what's in the data, at scale, across the estate and in detail before it is moved to the cloud. It's only once the data is discovered that it can be protected, deleted, and remediated.

A big mistake that cloud adopters sometimes neglect is security. We're often told by vendors that the cloud is inherently more secure than on-premise, and so organizations assume that in moving to Office 365 for example, Microsoft takes responsibility for handling all security risks. This isn't true. While cloud providers handle major risks quite well, they can't possibly address vulnerabilities that are unique to individual organizations. Therefore, the best way to keep data safe during the migration process is to ensure it is solid before it's moved.

Cleanse and classify data

Softwerx believes migrating everything is a bad idea because the cost of maintaining that data, backing it up and keeping on top of what's inside the data is becoming more difficult. And then there's the additional challenge of classifying data before it's moved to the cloud, which is crucial if the organization is to keep on top of data policies and processes once it is there.

> 84% of CEOs are concerned about the quality of the data they're basing decisions on, fearing it could lead to missed opportunities, lost revenue and reputational damage.[4]

Knowing what data the organization has and only migrating what's needed is powerful, not just because it saves money on storage costs by reducing the volume of data, but because the business gains confidence in the quality of the data that's migrated.

So how can organizations know what is actually lurking within the data estate?

The hidden depths of data

According to the Veritas Databerg Report, 81% of organizational data is duplicated, redundant, obsolete or trivial (DROT)[5], which means it doesn't need migrating to the cloud. Worse still, to migrate it to the cloud actually places the organization at greater risk.

Our own research and knowledge gained over many years in data discovery, tells us that every organization holds approximately 10GB of unstructured DROT data per employee. Typically, 42% of this data is company sensitive, and therefore should be encrypted, protected or deleted before a cloud migration. And 1% contains passwords in plain text – for example, where encryption keys are stored in plain text alongside databases.

The Marriott breach is a prime example of why organizations don't want DROT data in the cloud. Hackers found the encryption keys held alongside the 8 million encrypted credit card numbers that were stolen, as well as 5 million unencrypted passport numbers and personal details.

But how do organizations identify and distinguish the data that's usable and business critical from the data that needs deleting?

10GB of unstructured data per employee

Risk

Redundant Data on Decommisioned Servers

HR Database

Data Shared with 3rd Party

File Shares

Billing

Payments

Expired Data

Emailed Data

CRM

Structured Data   Unstructured Data

1%
of documents contain passwords

9%
contains personally identifying information

42%
Is company sensitive

46%
Is duplicated information

Data discovery: the enabler of successful data migration

200% greater ROI

"When you consider the laser precision that Exonar gives us to discover the data we need, and the peace of mind to delete with confidence, over a 3-4-year period that equates to an ROI of 200%."

John Harris, CTO, Mundipharma

The fact is, no business can embark on a data migration until they get a handle on all of their data – both structured and unstructured.

In the information age, organizations produce extraordinary amounts of data, which continues to grow in size every day. And while some of this data is neatly structured in protected systems and databases, the majority of it is lying unstructured across the estate, potentially unprotected against a breach.

The first step is to discover exactly what's in the estate. With full visibility over its data, an organization knows what data it has, where it is and who has access to it, so appropriate remediation actions can be taken to:

- **Secure data:** keeping sensitive, confidential and personal information secure.

- **Extract value:** mining intellectual property, research or insight within information and data.

- **Manage complexity:** managing the volume, such as migrating data to the cloud.

With Exonar, we've focused on how to provide unprecedented visibility of organisational data at scale, coverage and detail.

Coverage

Connect to all the most relevant data stores across your estate, no matter where or what it is, to get a truly organization-wide view.

Scale

Exonar Reveal creates an index of billions of items of your structured and unstructured data so you can search everything instantly.

Detail

Drill down to find the sensitive content of every item of information including the text in documents, not just the metadata.

We do this by enabling our technology to index any data set within any type of data repository, so it becomes instantly searchable.

Information indexing and classification

Organizations can't start to protect something until they know it exists. Only Exonar Reveal has the ability to identify ALL data sets – both structured and unstructured across cloud and on-premise – allowing users to query the data in any way they choose and produce real-time search results.

Think of Exonar Reveal like a reference book. Rather than start at the beginning every time to search for a precise piece of information (assuming the user even knows what this piece of information is), Exonar unlocks the power of the index. Now every piece of data is identified and recorded, with the technology constantly re-indexing the data to identify what's changed, what's new and to understand the context of the data.

Content and meta descriptions are added to enrich the index. And then it's augmented through:

▶ **Pattern matching:** resolves number and text patterns to identify information, like credit card numbers, government IDs, or phone numbers.

▶ **Natural language processing:** understanding and extracting the context of words, such as names, ethnicity and places.

▶ **Machine learning:** identifies or resolves additional contexts, such as topics and types of document like contracts or resumes and CVs.

> Exonar Reveal's indexing capability, along with the scale it can achieve with hundreds of billions of items searched across structured and unstructured data, is what makes the technology unique in the data discovery space.

Exonar Reveal is the layer in the technology stack that allows organizations to get the most out of their enforcement technologies. Because Exonar Reveal is focused on discovering data at scale, it's specifically built to plug into third-party systems to carry out the appropriate remediation actions, so users retain full control of their data estate and reduce the security threat.

> ## Things to think about....
>
> ▶ Do you know what proportion of your data estate is unstructured data?
>
> ▶ Can you identify risky data in your estate right now?
>
> ▶ If you have a data discovery tool in place already, does it allow you to index and show everything in your data so it can be cleansed and classified prior to migration?

Case study: protecting sensitive data prior to migration

An Exonar customer was decommissioning its legacy file share and email system to migrate to the cloud. The customer knew that there would be certain types of information, such as PII, stored in its unstructured data that was almost impossible to locate manually across the whole estate. However, it would need to be appropriately secured or managed before it was moved to the cloud.

The problem was that PII wasn't clearly marked or stored separately from the data that was fit to be moved wholesale to the cloud.

The customer asked us to help them find specific information across multiple storage repositories. Using Exonar Reveal, we indexed their data at scale to provide a single view of their data from one dashboard. At-a-glance the customer could see what they had, and if needed, drill into the data – down to specific documents – to create bespoke search queries.

The DROT data that was previously 'dark', lying lost and forgotten in shared documents, spreadsheets, emails and personal drives, was instantly revealed. With visibility of their data, the company could take action, which involved moving and deleting data at scale, optimizing storage requirements and placing the sensitive data in specific, highly secure locations.

With the peace of mind that its sensitive data was now secure, our client could finally embark on a cloud migration that enabled its digital transformation.

Data has strategic value

As a panellist at one of our recent round tables stated, "data is the new cash". Tackle the DROT data to cleanse it, classify the remaining documents and complete the migration, and organizations can start to ask the strategic questions that drive the business forward.

▸ **How can we get more out of the data by connecting data sets?**

▸ **How can we exploit that information to up/cross sell products?**

▸ **How can we commercialize data to extract more value?**

This is the future of data and information management. With data insight, businesses can make better decisions, mitigate risks and deepen the bonds of trust with customers. Once the business can get to its valuable data, it can mine intellectual property and use that insight to drive competitive advantage.

> ## ⟩ Exercise
>
> Ask: "If we slice and dice the information differently, what new use cases can we build inside the business to exploit the value of data?"

Ensure your migration is a success

Taking a data discovery approach to cloud migration enables an organization to ensure good data cleansing and management practices before moving the data estate. This enables the business to:

▸ Secure data: surfacing sensitive, confidential and personal information so it can be programmatically kept private and secure.

▸ Extract value: revealing intellectual property, research or insight within the information so it can be used to drive competitive advantage.

▸ Manage complexity: identifying the volume and complexity of data at scale, so it can be organized and migrated intelligently.

Should you need us...

The Exonar smart index enables data to be used to protect and power the organisation - billions of items all viewable and manageable in one place, and instantly searchable via a user-friendly browser interface.

To get you started, we're offering you the opportunity to see how Exonar Reveal can uncover what's in your data estate at scale, so you can see what you could do with that data to protect and power your organization. Talk to us about how we can help you.

Email: sales@anacomp.com

> "Exonar's software is the most comprehensive and effective data discovery solution we've been able to find."
>
> **John Harris, CTO**
> **Mundipharma**

References

1. https://www.techrepublic.com/article/73-of-cloud-migrations-take-a-year-or-longer-report-says/

2. https://www.panorama-consulting.com/erp-data-migration-and-cleansing-tips/

3. https://searchcloudcomputing.techtarget.com/definition/cloud-migration

4. https://www.forbes.com/sites/forbesinsights/2017/06/05/the-importance-of-data-quality-good-bad-or-ugly/?sh=626f38ce10c4

5. https://www.veritas.com/news-releases/2018-10-14-81-percent-of-data-stored-by-uae-organisations-is-unclassified-despite-improvements-in-data-management

Further reading

Blog: 3 transformational steps you need to take with data to drive a successful cloud migration

Report: The great data management conflict

Webinar: Data Discovery in Microsoft-centric Enterprises

Live demo: See Exonar Reveal in action

About us

Anacomp, Inc. a leading provider of technology-based records and data governance solutions, has partnered with Exonar Limited to deliver intelligent data discovery services to the U.S. federal government agencies and commercial organizations. Exonar is the world-leading data discovery company building the crucial software organizations use to discover their data at scale, keep it safe and realize its value.

All organizations face both the challenge and opportunity of owning large quantities of data and implementing a strong data management and governance strategy. Exonar is the only software available today able to find and reveal billions of items of structured and unstructured information; enabling businesses to 'know the truth' about their data in order to reduce risk, uncover business value and deliver digital transformation. By combining Anacomp's specialized expertise and experience implementing records management solutions with Exonar's world-leading data discovery software we can capture, classify and index large data estates efficiently. Exonar boasts a growing list of large and well-recognized customers in the pharmaceutical, financial services and national critical infrastructure industries.

Backed and funded by Beringea, Amadeus Capital Partners, Winton Ventures and Downing Ventures, Exonar boasts a growing list of large and well-recognized customers in the pharmaceutical, financial services and national critical infrastructure industries.

To find out more, visit **www.exonar.com**