Trial master file (TMF) processing using hyperautomation and digital content management

Atos

One of the most common truisms in every industry today is that “data is the new gold.” However, in the face of COVID-19, the most relevant word for the life sciences industry may be “faster.” Every sponsor company wants to beat the others by delivering vaccines or drugs to market more quickly, and all of them are sitting on data mines. Still, the traditional drug and vaccine development processes remain lengthy and very costly. Where do we go next?

Contents

Challenges and drivers to TMF operations

Trial master file (TMF) management is a clinical research and development process that typically involves processing huge volumes of document data manually, an inefficient approach that can cause reporting issues and compromise the quality of a sponsor's TMF.

In an age where buzzwords like digital transformation, hyperautomation and human intelligence are in vogue, and smart homes and smart cars are a reality, why can't we have smart TMFs?

Why can't all those cumbersome processes be automated — from filing, processing and quality checks to audits and submissions? From our perspective, that remains in the future, but for now, we will explore what is achievable in TMF processing using hyperautomation and content management.

Below, we will examine the challenges that the life sciences industry faces in eTMF (electronic TMF) processing, beginning with data preparation to archival and retention of data. Then, we will explore how content management technology, robotic process automation (RPA) and artificial intelligence (AI) can help overcome these challenges.

Challenge #1: An incomplete or fragmented view of the contemporaneous eTMF

Until now, most TMF processing systems have worked in silos. There are multiple stakeholders involved, including contract research organizations (CROs), labs, sites and vendors. A lack of accountability with multiple systems and multiple methodologies can lead to data gaps.

Sponsor companies must have complete insight into TMF completeness, timeliness and quality, because this insight can enable smooth inspections and submissions, as well as data archival for future use.

Challenge #2: Manual, monotonous, time-consuming TMF processing tasks

Ensuring TMF documents are fit for purpose requires a high degree of manual intervention to enrich the clinical data. Teams of knowledge workers must perform batch processing on documents and content, including the challenging process of manually classifying documents by type, extracting the correct metadata, and mapping it correctly based on the TMF reference model.

In the end, the finalized data must be interoperable following standard metadata, structure, content and naming conventions.

Challenge #3: Duplicate trial data document processing

Considering the enormous volume of electronic data generated before and during a trial, as a good clinical practice (GCP) it should be commonplace to maintain the clinical data applicable to multiple trials in the document repository. Moreover, any system should be expected to be intelligent enough to highlight the document data already present in the system.

There must be a provision in the system that enables knowledge workers to make the right decisions to categorize such data before it becomes part of the content repository for use during a clinical study.

Challenge #4: Growing volumes of historical clinical trial data eating up resources

In clinical research, massive amounts of data are generated from the inception to the end of a study. It is important to maintain the right balance of data storage between active systems and systems intended to manage document data during a clinical study or pre-study. Once trials are completed, data must be managed or preserved in a way that does not create resource or performance issues.

The role of hyperautomation in content management

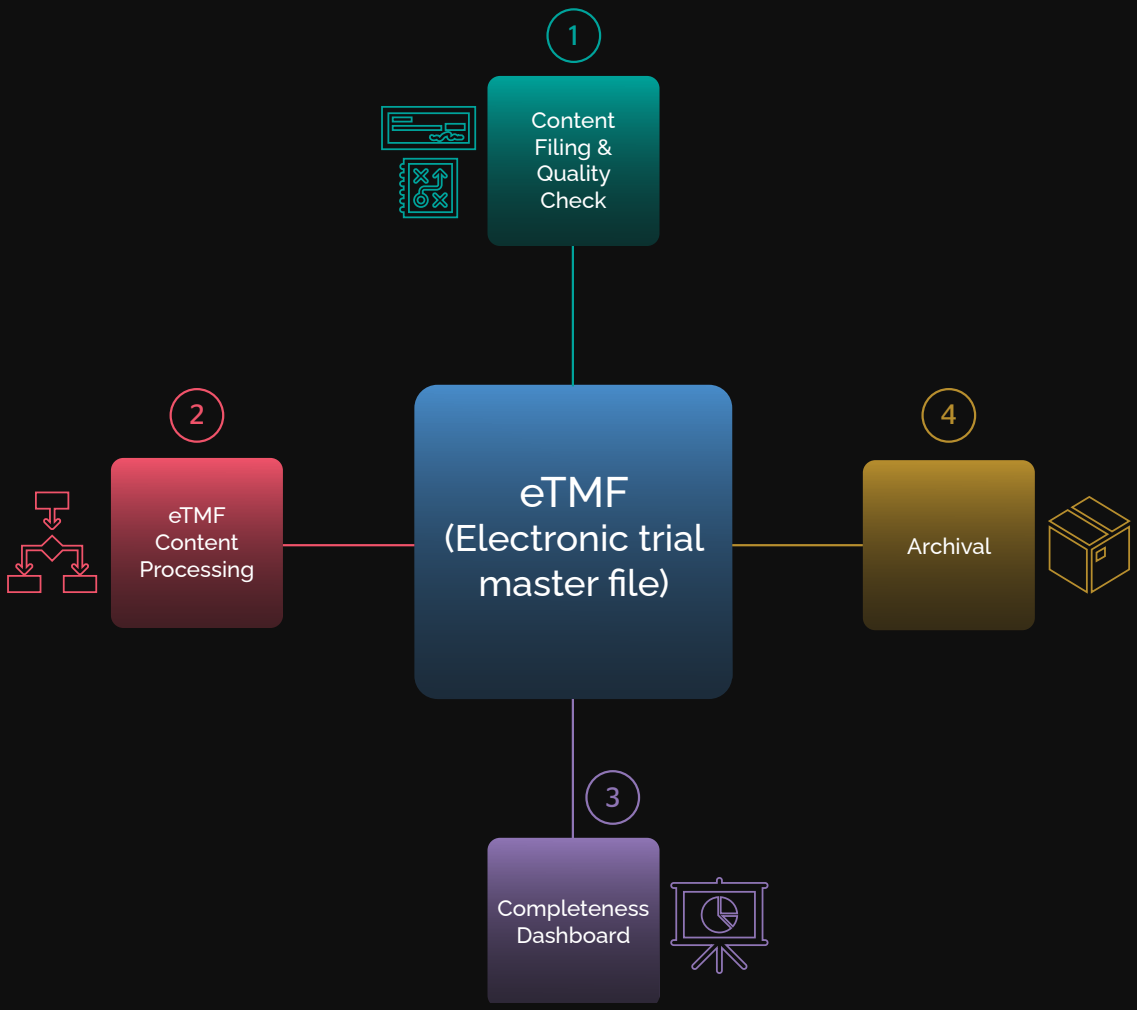
Employing hyperautomation with content management is a solution with the potential to resolve all the challenges outlined above.

Hyperautomation can provide a structured and efficient approach to eTMF processing — enabling content to be immutable and archived based on retention policies and compliance rules, encompassing:

- Automated TMF document batch processing
- Classification
- Filing
- Document data processing
- eTMF workflows and lifecycles
- BI reporting

Advanced content management solutions provide seamless connectivity and collaboration between contract research organizations (CROs) and sponsors, ensuring the eTMF serves as a single source of truth for TMF document data.

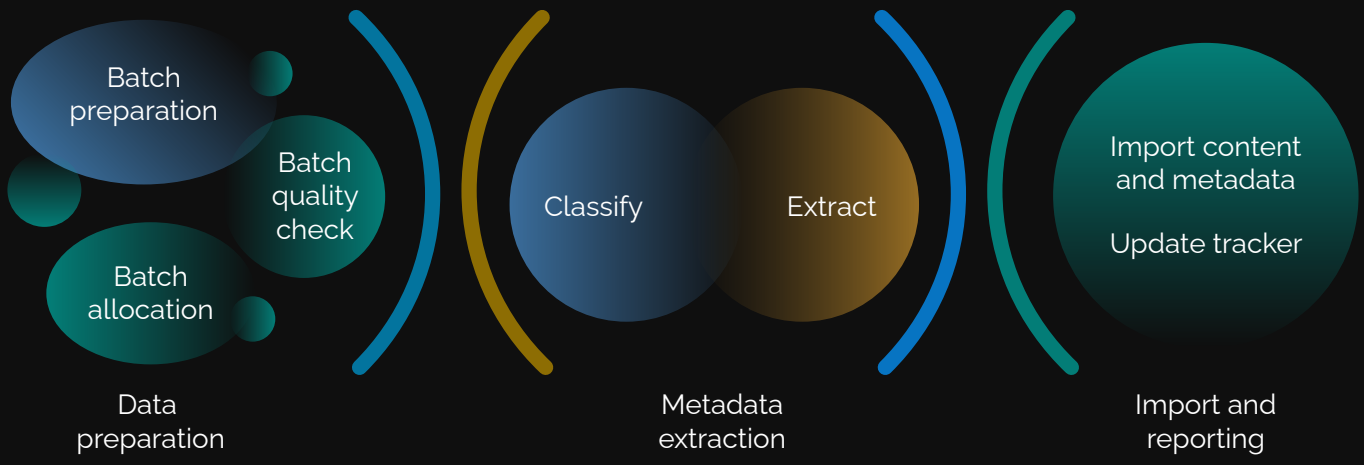
However, any hyperautomation solution capable of addressing the major challenges must consist of the following four modules:



1. Automated content filing and quality checks

Making documents fit for purpose and mapping them to a TMF reference model is the first key step to automated content filing in eTMF processing. Even before COVID-19, sponsors were required to process huge volumes of semi-structured and unstructured data during a clinical study — and it is only increasing.

The three key components of the automated content filing and quality check process are:



Batch data preparation

- RPA-enabled automatic download of study/batches
- Tracker (Excel) preparation with data
- Duplicate batch tracker check
- Batch quality checks

Document classification and metadata extraction

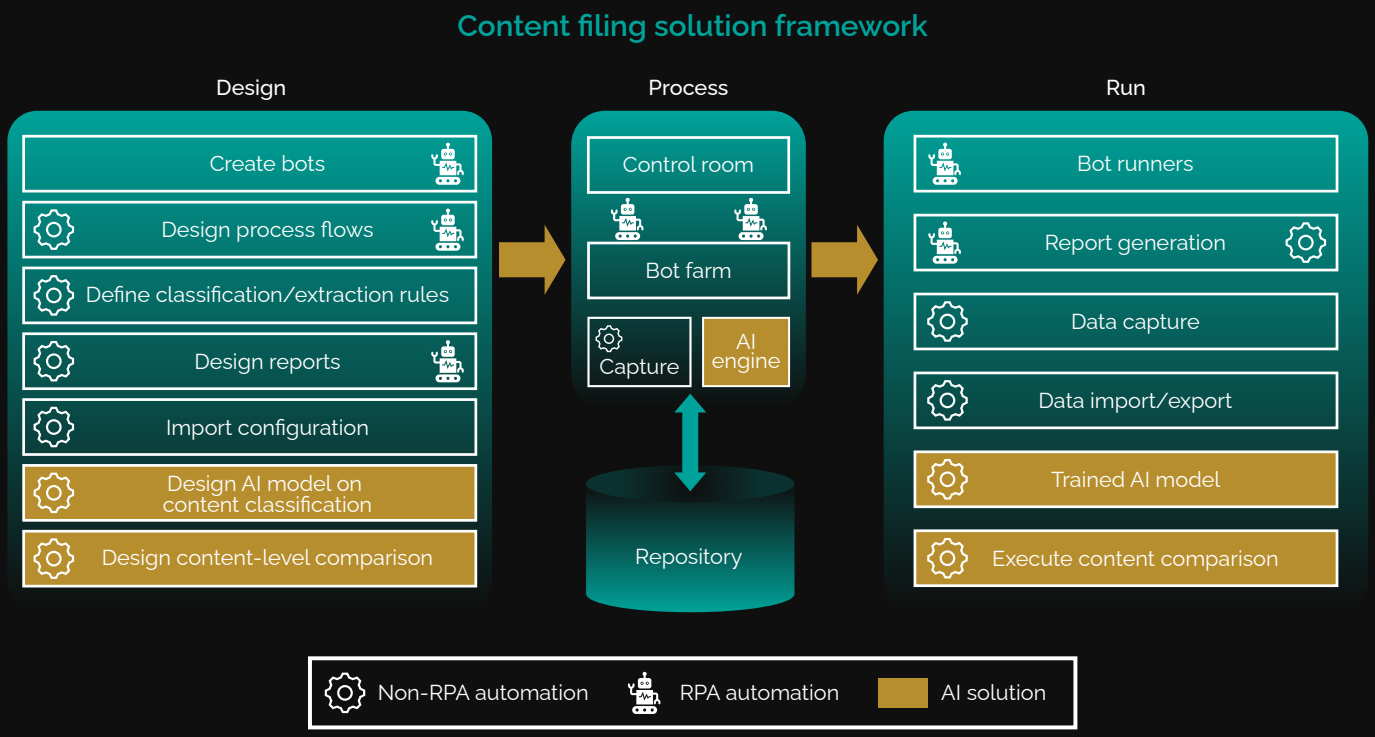
- Train ML algorithm to automatically differentiate and classify documents
- Identify document type based on trained AI model, OCR recommendation
- Remove blank pages and redact content as per HIPAA requirements
- Automated wet signature identification through capture solution
- Compare clinical data forms based on metadata and content

Import and reporting

- Import documents to repository using custom API or out-of-the-box connector
- Automatically update clinical data content trackers, to enable clinical team to monitor clinical study batch data filing
- Communicate overall status to clinical team via customer portal / email to share success and failure scenarios for eTMF document data processing

In this phase, the goal is to save time and efficiently file all required clinical study content with better quality and with low overhead. Hyperautomation and content management enables eTMF document data to be preserved in a standard format and structure — with appropriate metadata and naming conventions for easier retrieval and more effective dashboards and reporting.

Below, we have depicted a solution framework that employs intelligent capture, RPA bots and AI to automate the content filing process.



All the RPA and intelligent capture processes are designed in the design module and executed in the run module. The process module is responsible for managing the running instance. It is designed to satisfy quality checks and is responsible for converting eTMF data in TMF compliance mode for further processing or the next stage of eTMF processing.

2. eTMF processing using a TMF-compliant content management solution

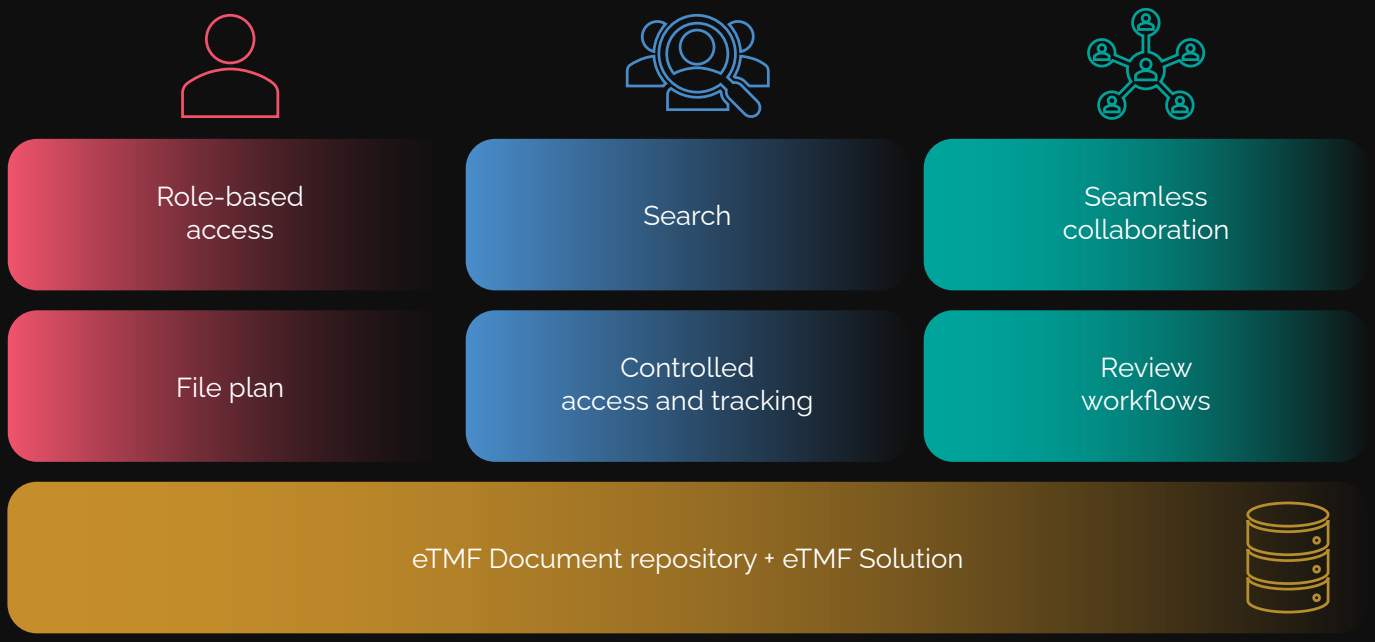
This is the core phase of TMF processing, where sponsors and CROs use a secure system to collaborate, preserve and manage the trial data. Broadly speaking, this phase consists of:

- **Automated file planning at a product, trial, country, and site level**
  The moment data becomes part of document repository, it is required to be preserved in a TMF-compliant format to be retrievable for audit and submission.
- **Managing file plan with forms**
  The file plan must be associated with product registration, trial registration, country registration and site registration.
- **Creating trial documents and controlling access** to document data with required document tracking
- **Reviewing and approving trial document workflows**
- **Locating and finding the required documents** based on site, product, clinical trial, and document type

A **content management solution** will help both before trial begins (by accelerating startup) and during the trial (by improving the study's overall transparency). All data must be maintained according to Drug Industry Association (DIA) guidelines.

For clinical trials, OpenText™ Documentum™ for eTMF or Lifesciences Document Management helps effectively plan, collect, track, and maintain essential good clinical practice (GCP) compliant clinical trial documentation.

To preserve and manage the eTMF data captured according to the TMF reference model requires an advanced content management solution framework like the one depicted on the next page.



Once data are processed by the content filing solution, eTMF data must be preserved and processed to be used throughout the clinical study over a period of approximately five to seven years. Accordingly, the eTMF content management solution should have the basic components shown above. In addition, the data model must comply with the TMF reference model — whether custom built or a commercial application like OpenText Documentum eTMF or Veeva Vault.

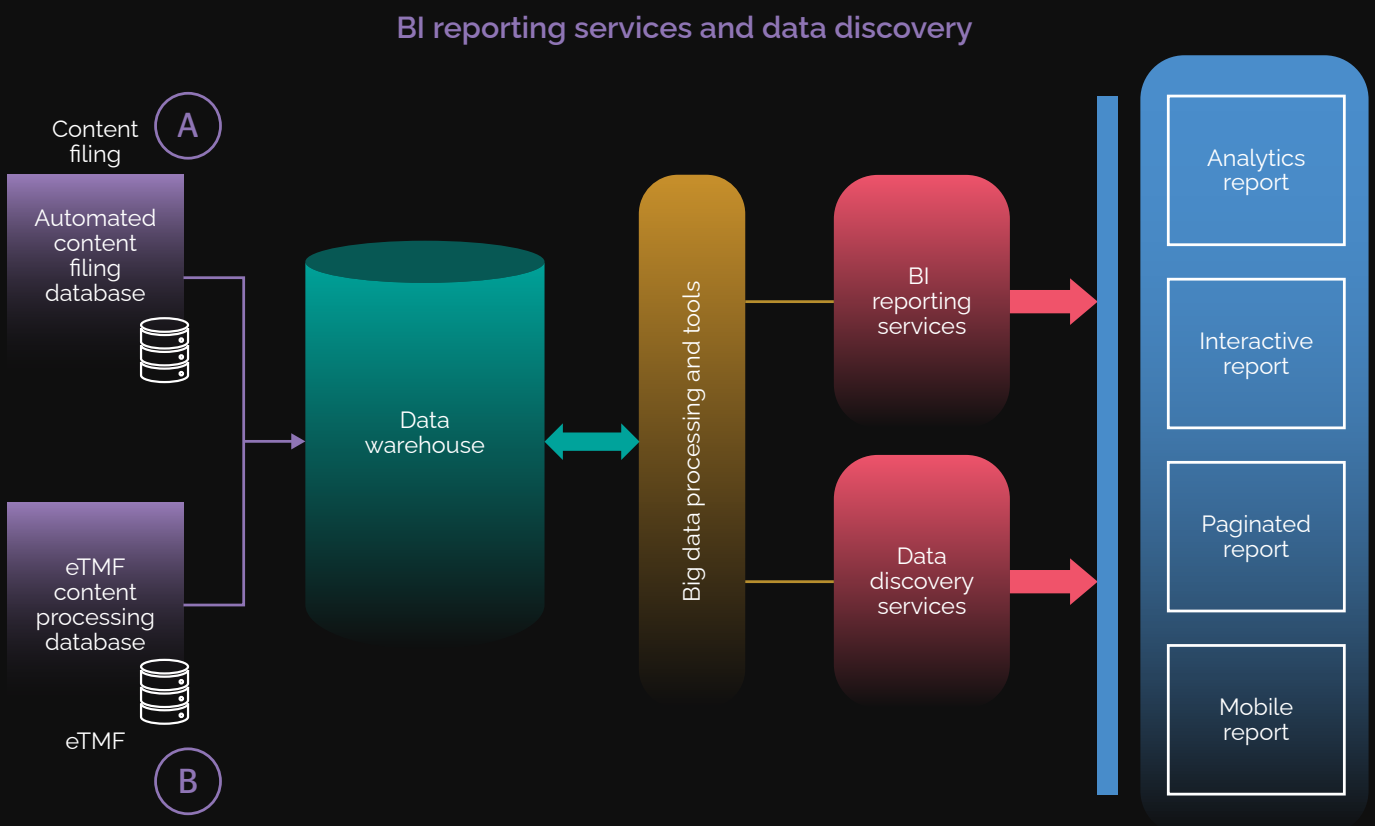
3. BI reporting and interactive TMF completeness dashboards

Throughout the TMF processing journey, the real challenge is to keep up with ever-increasing volumes of documentation. Typically, clinical teams maintain and track records on Microsoft Excel sheets, using built-in visualization tools to verify that documents are being collected and completed contemporaneously.

However, achieving a deeper level of insight into TMF processing requires adding a trial status dashboard on top of the content management solution, enabling you to understand the current state of a study based on site, country and all other required KPIs. Below are some important criteria you should aim for or expect from a TMF report or dashboard:

- Consistency with TMF reference model
- Track study milestones as per the TMF plan
- Track TMF completeness according to the expected document list
- User-friendly interface that supports finding information easily
- Audit details on study with meaningful statuses
- Robust security to enable a preview of document data with required data protection

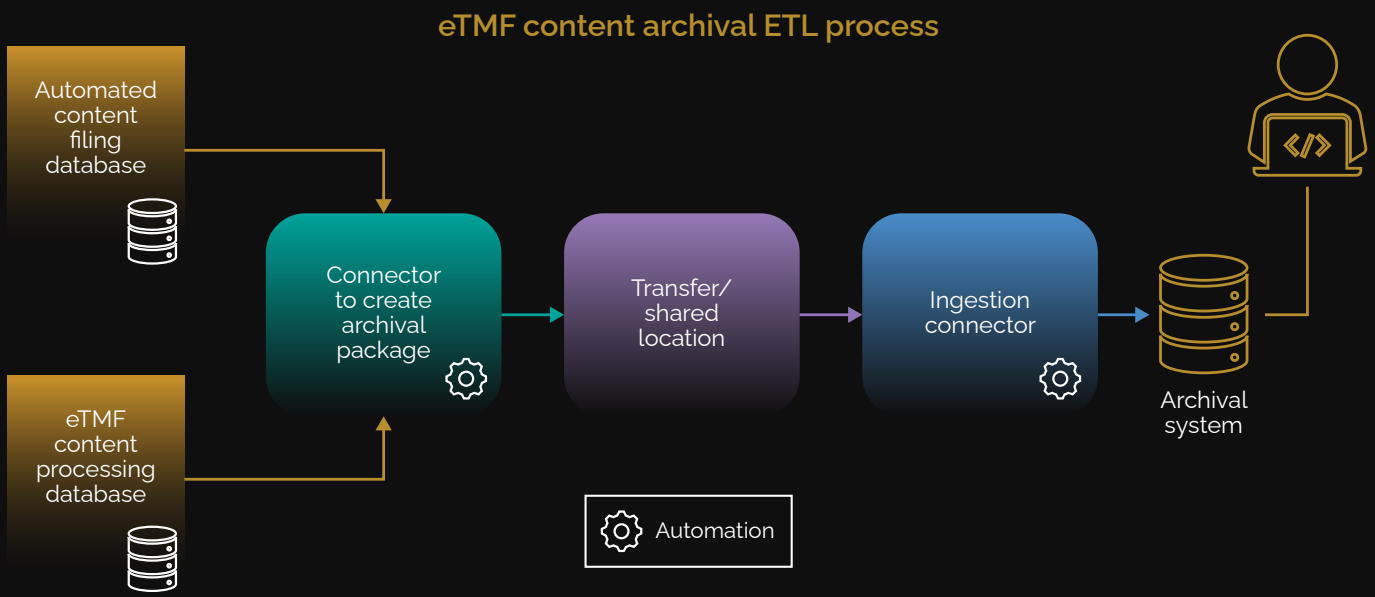
BI reporting services can be employed to create a completeness dashboard to satisfy clinical team KPIs (as mentioned above) or to track completeness. We recommend combining content filing data
A. and eTMF data
B. to be transformed in a data warehouse. In addition, strong analytics must be layered on top of the data warehouse in order to slice and dice the data to serve business purposes. A BI reporting solution can be used to obtain the required output in the form of analytical, interactive, paginated and mobile reports.



4. Clinical content archival

After a certain period, the data generated or converted into electronic format during processing must be archived as historical data. It is important to ensure that this data can be made available for clinical study reference whenever required.

The archival process helps ensure that the existing eTMF document processing system does not suffer any performance impact, and that data can easily be archived from the existing system to the archival system.
The most important feature of archiving is governance and compliance with health authority retention policies, to avoid penalties and make historical data available for future clinical trials.



The purpose of archival is to manage and store immutable data in a compressed format and make data available whenever is it required. To achieve this, a specific set of connectors must be written to extract, transform and load (ETL) the data in the required format for the archival system. The central transformation step is the most critical part of this process.

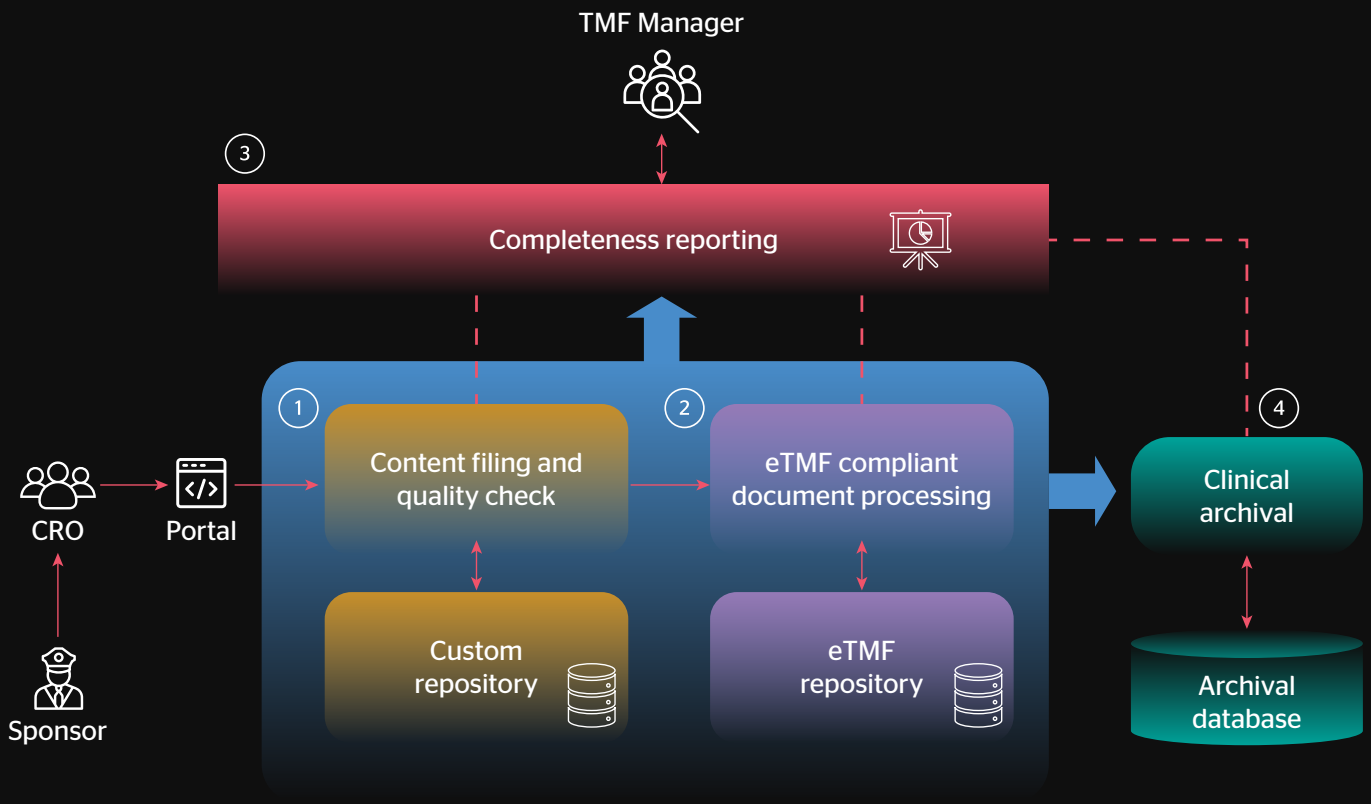
Atos works with clients to perform the required transformation of data to be migrated into the archival system. Using web services, the data can be made available on-demand to any third-party system.

The Atos approach

Based on our experience, we believe that hyperautomation with digital content management technology can solve most, if not all the problems that life sciences companies face today.

They have the potential to bridge the gaps between CROs and the sponsor's clinical trial team, enabling efficiency, transparency, and timeliness in establishing a truly contemporaneous eTMF. Furthermore, we anticipate that hyperautomation with content management has the potential to enhance the time to market and enable life sciences companies to increase competitiveness.

The Atos approach to TMF hyperautomation with advanced content management has been built considering all the best practices we have outlined earlier, and includes the four major modules shown below.



However, before embarking on any hyperautomation journey, there are a few important considerations that must be considered:

- The ROI from AI technology is not immediate – you may need to wait to see the outcomes anticipated
- It's important to decide up-front whether to follow a one-step approach or to scale gradually
- Hyperautomation with content management is not just a tactical move for short-term advantage, but a strategic long-term differentiator
- Adopt a top-down approach to bring automation and intelligence
- Employ a "train the trainer" approach to educate your existing knowledge workers

Conclusion

There are many other use cases for hyperautomation technology, such as using a retrospective analysis of historic TMF data to make faster decisions about site selection or initiation based on a site's historical performance. This type of analysis helps provide insights about high-risk areas for TMF from a quality and completeness perspective, which can help implement a risk-based quality methodology around TMF or reallocate resources to critical areas.

Because all the data generated from a TMF is part of a larger digital ecosystem used by multiple systems like clinical trial management systems (CTMS), regulatory and safety to collaborate, employing hyperautomation to produce uniform data will lead to smoother, more accurate submissions and drive the ultimate goal of bringing drugs to market faster.

Learn more about Atos life sciences solutions at atos.net/en/lp/life-sciences

About Atos

Atos is a global leader in digital transformation with 105,000 employees and annual revenue of over € 11 billion.

European number one in cybersecurity, cloud and high performance computing, the Group provides tailored end-to-end solutions for all industries in 71 countries. A pioneer in decarbonization services and products, Atos is committed to a secure and decarbonized digital for its clients. Atos operates under the brands Atos and Atos|Syntel. Atos is a SE (Societas Europaea), listed on the CAC40 Paris stock index.

The purpose of Atos is to help design the future of the information space. Its expertise and services support the development of knowledge, education and research in a multicultural approach and contribute to the development of scientific and technological excellence. Across the world, the Group enables its customers and employees, and members of societies at large to live, work and develop sustainably, in a safe and secure information space.

Find out more about us
atos.net
atos.net/career

Let's start a discussion together

For more information:

Rachana Dubey
Sr. Consultant H&LS Practice Portfolio
Linked in Profile

Mitesh V. Kothari
Principal content architect - ECM
Linked in Profile

CT-210629-EM-WP-eTMF