Article

Essential Steps for Intelligent Document Processing in Clinical Trials

AIThority: Automating clinical trial processes using AI

The healthcare industry is undergoing a digital transformation, driven by the adoption of artificial intelligence (AI) and machine learning (ML) technologies. These technologies are being used to automate and streamline processes and improve data management, helping to accelerate the development of new treatments and therapies, improving patient outcomes and, ultimately, saving lives.

Clinical trials generate vast amounts of data across various formats, including text, audio, video and images, which can be difficult to collect, organize and analyze, and can also slow down the development of new treatments and therapies. Intelligent document processing (IDP) is one technology being integrated into clinical trials to address this challenge. Applying AI and ML techniques to IDP to process structured, semi-structured and unstructured documents to enable systems and applications to read and process content in documents in a similar way a human could.

IDP helps to automate the collection, organization and analysis of clinical trial data. Through this process a team can simulate and predict trial outcomes, helping researchers uncover hidden patterns and trends in clinical trial data. This certainly can lead to new breakthroughs in medicine.

IDP Implementation in clinical trials

Given the number of clinical trials taking place, the adoption of IDP can deliver significant improvements, boosting productivity and accuracy, saving time and money and streamlining the clinical trial process. However, as the saying goes, 'fail to plan, plan to fail,' and implementing IDP in clinical trial processes is no exception. Sponsors must start by carefully considering their goals and objectives, understanding their document processing requirements, identifying the types of documents that need to be processed and pinpointing any areas for improvement.

It is crucial to be cognizant of the challenges that pharmaceutical companies encounter in their data flow during clinical trials. For instance, manually populating site folders and electronic trial master files (eTMFs) is not only time-consuming but can also lead to compromised document security and data privacy, archiving and retrieval difficulties and human error, adding a failure rate of up to 25%. Additionally, the stringent regulatory and security requirements mandate that IDP systems ensure patient privacy protection and the clinical trial to maintain audit trails while managing unauthorized access to sensitive data.

To overcome resistance to change in this traditionally conservative industry, it is essential to demonstrate the value of AI-driven solutions and ensure they are secure, private, and compliant. Choosing an IDP solution that caters to the organization's specific requirements, including support for relevant document formats and languages, is crucial. Assessing the solution's AI/ML capabilities, its ability to integrate with existing systems, incorporate technology and data science and integrate AI and generative AI is essential for handling digital content efficiently.

To successfully implement IDP in a clinical trial's digital content flow, companies must follow these essential steps:

1

Quality auto-review: Data quality is paramount, and since AI's effectiveness depends on the quality of its training data, it is critical to promptly identify incoming content and verify its layout correctness and data convertibility. Clean and organize the data to ensure it can be processed by the IDP solution. This involves tasks like removing errors, converting images and scans to text and standardizing data formats. If the system detects any issues, it should flag the information for user review and resubmission in real time, expediting data collection and quality control processes.

2

Digital twin and classification: Deploying digital twins facilitates digitization of all clinical trial content for universal accessibility. This data can be used to train generative AI models to recognize patterns and relationships. Companies can also utilize digital twins for pre-audit checks, proactive eTMF and electronic common technical document (eCTD) population and glean insights from past trials.

3

Auto-translations: The ability to automatically convert content into other languages is crucial for deploying IDP. This can be achieved using domain-specific regulatory or safety language ontology sets. Auto-translation streamlines communications and drives efficiency gains.

4

Sensitive data handling: Data privacy is an integral part of any clinical trial, necessitating the automation of safe processes for sourcing, linking, combining, reusing and sharing protected data. Deploying privacy analytics enhances the retention of sensitive data by implementing redaction capabilities and providing only relevant content to users.

5

Entity extraction: Once content is digitized, teams need to identify the sections within that content and locate specific information. NLP and natural language understanding (NLU) enable comprehension of text and its meaning. For instance, NLP can be used to analyze scheduled assessments for a patient and identify their participation requirements. This information can then be utilized to develop appropriate models for optimizing patient burden management.

6

Insights and best actions: Technology can assist in content analysis to generate actionable insights in areas such as risk assessment, patient burden, protocol amendments, potential outcomes and theoretical modeling. The convergence of digital twins and NLP enhances data understanding, enabling generative AI models to learn essential patterns and relationships for precise predictions and creative content creation.

Automating IDP brings numerous benefits

Automating IDP is gaining traction as a solution to legacy IDP challenges, including expediting operations, enabling continuous processing, enhancing accuracy, fostering collaboration, and ensuring regulatory compliance.

A significant advantage is the ability to promptly assess the quality of content from trial sites before its integration into the eTMF for final trial file storage. This is achieved by implementing an API-enabled SaaS solution for automated IDP within site folders. This process allows for early identification and resolution of issues such as missing signatures, scan issues, and layout problems.

Automated eTMF systems address manual processing challenges by offering features like document version control, audit trails, notifications, remote accessibility and advanced search capabilities. They eliminate manual eTMF entry while maintaining quality. These solutions index documents, automate workflows, assist translations, reduce processing time and free up employees for value-added tasks. Additionally, they can handle scans and images in any language, extracting metadata and creating digital twins for improved recognition.

Another key benefit is the ability to gain clear insights into whether trial sites have reviewed, acknowledged and understood protocol amendments. This is crucial for optimizing communications with sites and ensuring regulatory compliance. Automated IDP solutions provide exceptional visibility through version control and audit trail monitoring.

IDP's Future role

As pharmaceutical companies strive to automate and streamline clinical trial processes, IDP's are gaining prominence, driving greater efficiencies and unlocking deeper research insights. Companies are increasingly exploring generative AI applications for data mining, template creation, quality control, site communication and clinical trial operation guides.

AI's ability to rapidly identify potential trial participants from medical records and promptly analyze medical data to detect safety issues holds immense potential. In the near future, we anticipate pharmaceutical companies will establish a foundation for their long-term AI roadmap by developing and deploying "mini" versions of generative AI models internally. This will allow them to harness the benefits of AI while safeguarding the quality and security of their sensitive data. In this context, IDP serves as a powerful tool for organizations to enhance the efficiency and effectiveness of their clinical trials. Following the steps outlined above, organizations can successfully implement IDP and begin reaping its benefits.

About the author

GARY SHORTER
Head of Artificial Intelligence at IQVIA Technologies

Gary pursues the use of emerging technology to provide new and more efficient capabilities to enhance clinical trial management. This includes development of new design software through to more recent advancements with AI/ML capabilities where his team has developed several micro-products and micro-services that can be plugged in and used by any SaaS solution.

IQVIA TECHNOLOGIES

CONTACT US
OrchestrateYourTrials@iqvia.com
iqvia.com/CDAS