celegence

Lean Authoring

Eliminating Redundancies And Complexities For The Pharmaceutical Industry And Health Authorities

Checklist of lean authoring best practices for efficient data exchange

Table of Contents

Table of Tables

1. Introduction

A **Health Authority (HA)** must safeguard patients' health using medicinal products. To do that, HAs require documented proof of the quality of the product and its production process, the product's nonclinical pharmacology and toxicology data, and the clinical safety and efficacy.

The development costs of a medicine are comprised of the costs for pharmaceutical development, production of the pharmaceutical products, the nonclinical pharmacology and toxicology studies and the clinical studies. On top of that, proof of a medicinal product's quality, safety, and efficacy needs to be documented to allow health authorities to review and approve it for use in well-controlled clinical trials and eventually in a less controlled market situation.

The reality is that therapeutics need to be approved by each region or country separately; therefore, respective, tailored dossiers are required. If there is obvious documented proof, all health authorities will simultaneously come to the same conclusion. However, HAs may have their own requirements, various levels of expertise, and prior knowledge when evaluating the information in a dossier. As a result, some HAs might request more detailed information or apply narrower acceptance criteria than others. Even differences in the packaging (e.g., tablets packed in blisters versus bottles), different ranges of strengths, and various manufacturers used to supply a market may result in the need for diverging information to support a marketing authorization. For marketing authorization holders, this results in complexity in tracking what version/configuration of a medicinal product has been submitted to and approved by the respective HAs. From their perspective, HAs aim to objectively review and align their assessments of similar or identical applications for marketing authorizations.

Currently, the main document format to exchange information between applicants and HAs is PDF. Most of the information is shared in narratives, supported by graphs and tables. Sometimes, the information in the table is even reworded into a narrative. Why is data not presented as data? Are analyses run on the data, and only interpretations of the analyses are captured as narratives? Because it has been done like that since the introduction of paper dossiers over 80 years ago!

The US Food and Drug Administration (FDA) clearly recognized the shortcomings of data locked in PDFs and, over time, requested clinical data (and later also nonclinical data) in datasets. Other HAs only require entry of some key characteristics of a drug in a database in addition to the eCTD (e.g., in the European Union (EU), entry of key SmPC elements in XEVMPD, and recently implemented in an electronic application form). Lean authoring is a recognized approach to bridging the gap.

Lean authoring allows a reader to focus on the relevant data and prevents data from being duplicated in different documents. Every duplication has an inherent risk of inconsistencies. Lean authoring further facilitates document maintenance when addressing the impact of a proposed change in a drug's production, quality, or usage. It also decreases the number of documents to be managed and the time to author, review and maintain documents (see Table 1).

Table 1. Impacts Of Lean Authoring For Industry & Health Authorities

Topic impacted by Lean Authoring	Impact for the Pharmaceutical Industry	Impact for the Health Authority
Quality of information Number of documents in the initial MAA increases because of finer granularity of smaller documents. The total amount of content reduces	Decreases the total number of documents in the document management system, because many duplications of large documents to accommodate country-specific deviations are reduced.	The eCTD browser allows reviewer to see a finer granularity of documents as if it were chapters in a larger document.
Efficiency Number of content pieces to implement a change decrease because of the lack of repetition	Facilitates easier lifecycle management of the dossier to reflect the product lifecycle and a better reflection of where the changes are.	Less to review, because less documents are impacted by the easier lifecycle management of the dossier.
Global accessibility	Modular structure allows reuse of information that can be adapted easily for different markets and regulatory environments. Given the ability to view across dossiers, it will become transparent on what document has been used where and when to support a single change in the product or its usage.	More efficient health authority assessments through precise focus on the change and its impact on the appropriate documentation.
Cost savings, streamlined Total volume of content diminishes due to a lack of redundancy and repetitions	Less content to be authored, reviewed and released.	Less to review, because of the lack of redundancies.
Improved quality and agility Factual data in tables	Easier to collect relevant data and apply IDMP attribute names and values in accordance with EMA SPOR or US SPL. Ability to adapt faster to changing regulatory requirements (e.g. addition of providing a new data element).	Easier to retrieve relevant data because of the semantics applied and no narratives to absorb and extract data from. Subsequently, HAs can run their own analyses of the data and can do comparisons with other applications or inspections of R&D or manufacturing sites.
Scientific evaluations	Limited to narratives in dedicated sections in study reports and Pharmaceutical Development and Method and Process Development in Module 3. All evaluations should be based on the data provided in tables.	The narratives with the scientific evaluations can be compared to the own analyses of the same data. The scientific evaluations and judgement of claims are limited to a few documents and no longer scattered all over the Module 3 documents or even lacking.
Accelerated timelines	▪ Faster time to market for reduced drug development cycle time and maintenance time. Improves patient safety. ▪ Saves costs.	▪ Greater bandwidth to approve more therapeutics ▪ Allows for objective evaluations and less influence by subjectivity and experience level of the assessor. ▪ Reduces costs.
Improved relationships and trust **Focus on patient health and safety**	Pharmaceutical companies build greater trust and credibility with regulatory bodies through consistent compliance. Gain greater trust with patients by focusing on patient health.	Health authorities and notified bodies build better relationships with pharmaceutical through more predictable information, less revisions, and greater focus on patient health.
Audits simplified	Improved documentation helps the submission process and sets pharmaceutical companies up for success on audits and improves quality.	Saves time and audit costs through better preparation from pharmaceutical companies and quality documentation.

2. What Is Lean Authoring

Lean authoring aims at eliminating unnecessary words, punctuation, complex sentences etc. to make writing as clear and succinct as possible. This style of writing facilitates focus and makes use of simple words and short sentences. It is not only about what is written but also about how it is written and presented (visual style and overall readability).

Lean authoring involves streamlining the writing process and reviewing, publishing, and maintaining standalone reusable documents using standardized templates, modular content, and meaningful, unambiguous, relevant cross references.

The goal of lean authoring is:

- To speed up the knowledge transfer by reducing the verbiage for documented content
- To re-use documents rather than re-create
- To improve the reliability of the knowledge, leaving less uncertainty to the receiving party of the document
- To allow for faster decision-making.

The standardized structure of a Common Technical Document (CTD) for pharmaceutical products facilitates the submission of a regulatory application consisting of multiple documents, each containing specific information. The CTD consists of a coherent assembly of relevant documents. Lean authoring becomes important when applying a lifecycle to documents and dossiers, e.g., in the electronic CTD (eCTD). For example, replacement documents can only be applied to whole documents and not to part of the documents. In addition, the documents currently approved by the HAs must comply with the current production process, quality and proper usage of the medicine. Mere inconsistencies throughout the approved documents of a regulatory dossier represent non-compliance.

The principles used to author documents can significantly impact the efficiency and quality of the output. Traditional authoring often involves a linear, extensive approach with multiple drafts, lengthy processes, and extensive revisions. By contrast, lean authoring focuses on streamlining these processes by delivering value through iterative improvements and reducing waste. Lean authoring provides factual data in easy to consume formats such as tables and bulleted lists which organizes information in a clear, concise manner. Moving from traditional authoring to lean authoring involves adopting several key principles to enhance knowledge transfer, readability, and data consumption.

By integrating lean authoring into operations, the pharmaceutical industry can enhance its ability to produce high-quality, compliant, and impactful content efficiently. This will ultimately improve overall productivity and better serve patients and healthcare professionals.

A comprehensive comparison of lean authoring principles to traditional methods is provided in Table 2, later in this document.

The field of lean authoring and optimizing for authorizations will continue to evolve. The future will continue to focus on anticipated changes in regulatory requirements, preparing for future compliance challenges.

3. A Mindset Change Is Needed

3.1. Evolution of Regulatory Documents

Of course the pharmaceutical industry collects enormous amounts of data, they have the data in databases; they do their own analyses, write reports to evaluate the findings and make decisions to develop a drug further or to terminate the development. Similarly, all the disciplines write development reports, progress reports and study reports under the predicate rules of GLP, GMP and GCP. Though the study reports compliant with GLP and GCP can often be used as such in a regulatory dossier, the GMP reports are not submitted to the HAs. For Module 3, those GMP reports are rewritten to Module 3 documents.

Why not hand over the relevant data from industry to HAs in the first place? Why reuse data and rewrite information in PDF documents by industry that subsequently must be interpreted and verified by HAs? The answer lies in the data model, semantics, reliability of the data and exchange standard.

Example concepts to incorporate:

- Think about how the data will be used and consumed – create for consumption
- Prepare ease of data extraction and processing for regulatory authorities
- Simplification of data review and analysis
- Present data in a table and omit the conversion to prose (repetition)
- Avoid scanned documents - capture as an image

Publishing and document management in pharmaceutical companies have evolved from manual, paper-based methods to digital platforms and lean authoring practices, improving efficiency in regulatory document creation and distribution. This evolution has prioritized accessibility, compliance, and patient safety through streamlined processes and enhanced content management systems.

Specifically, the formats have evolved. Originally, submissions and documents were on paper. This moved to digitization with PDFs. However, PDFs did not solve the need for sections and data

to be sent to different places. The next phase was a database like the eCTD format. But despite all the advantages of eCTD, it is still a collection of PDFs. The change needs to be at the input level. The real transformation is about manageable elements that can be dynamically grouped to be sent with the right content and context.

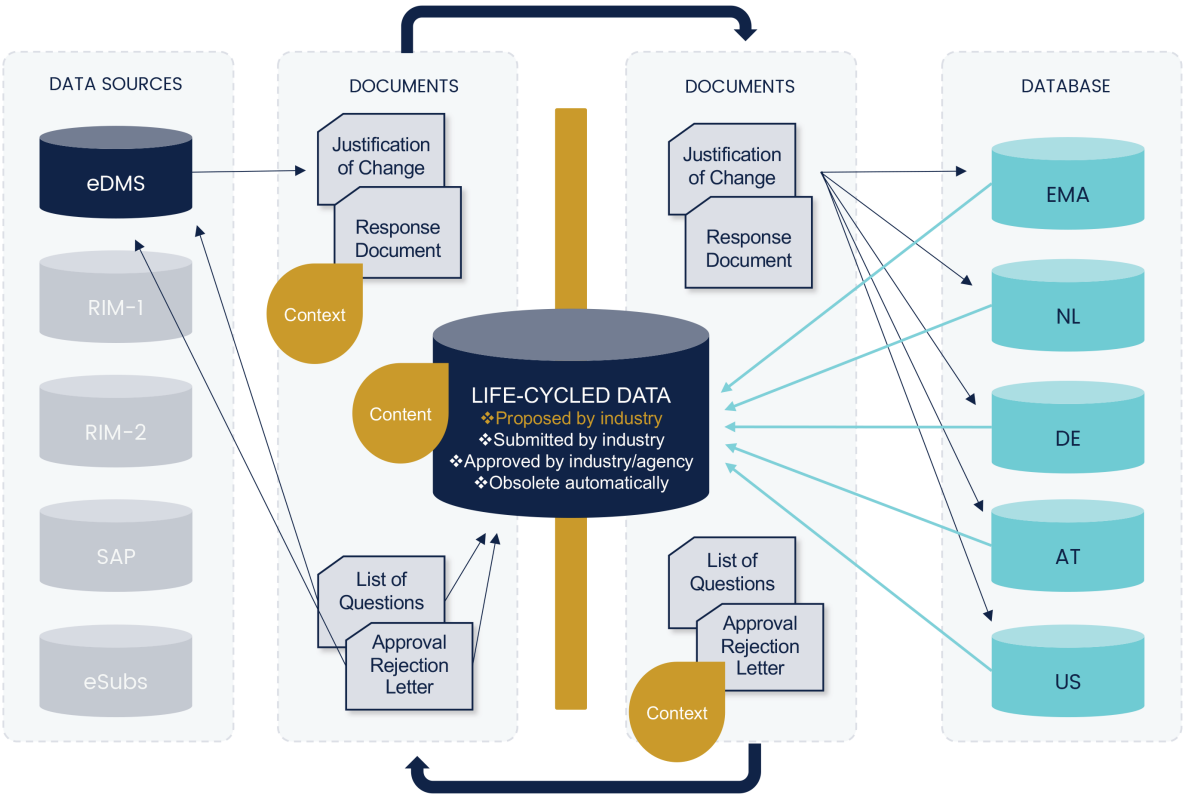
IDMP, PQ/CMC and ICH M4Q Revision 2 aim for more factual information in databases. Scientific evaluations can be limited to eCTD Module 2 and electronic product information that can immediately be linked to the scientific evaluations, which are linked to the analyses based on the data in the database. The exchange of information might change from sending packages to sharing the information in cloud information.

3.2. Future CTD Module 2.3 Quality Overall Summary and CTD Module 3

The future of lean authoring in ICH M4Q and beyond will likely be characterized by a greater reliance on technology, increased collaboration, and a focus on data-driven decision-making. These advancements will make the authoring process more efficient and improve the quality and compliance of regulatory submissions. As the industry evolves, companies adopting these practices will be better positioned to meet regulatory requirements and bring new drugs to market more swiftly and effectively.

Leveraging Lean Authoring Method – Output Centric:

Future data and document exchange within and across industry and agencies



Reference: "From CMC document to CMC data" meeting 10/10/17; Charles Morgan, Rodrigo Palacios, Hans van Bruggen

4. Principles Of Lean Authoring

4.1. Principle 1: Separate Text And Data

The first principle is the most important, separating out text and data. This sets up the data to be reused in multiple documents with the appropriate text 'wrapper' for context. The writer will want to avoid long, verbose prose or reaching a certain number of pages.

> *"Essays should be long enough to reach a meaningful destination and cover essential points, but short enough to maintain interest along the way."*
>
> **Hans van Bruggen, 2024**

Tabular data and text/prose are complementary tools, each for the right job:

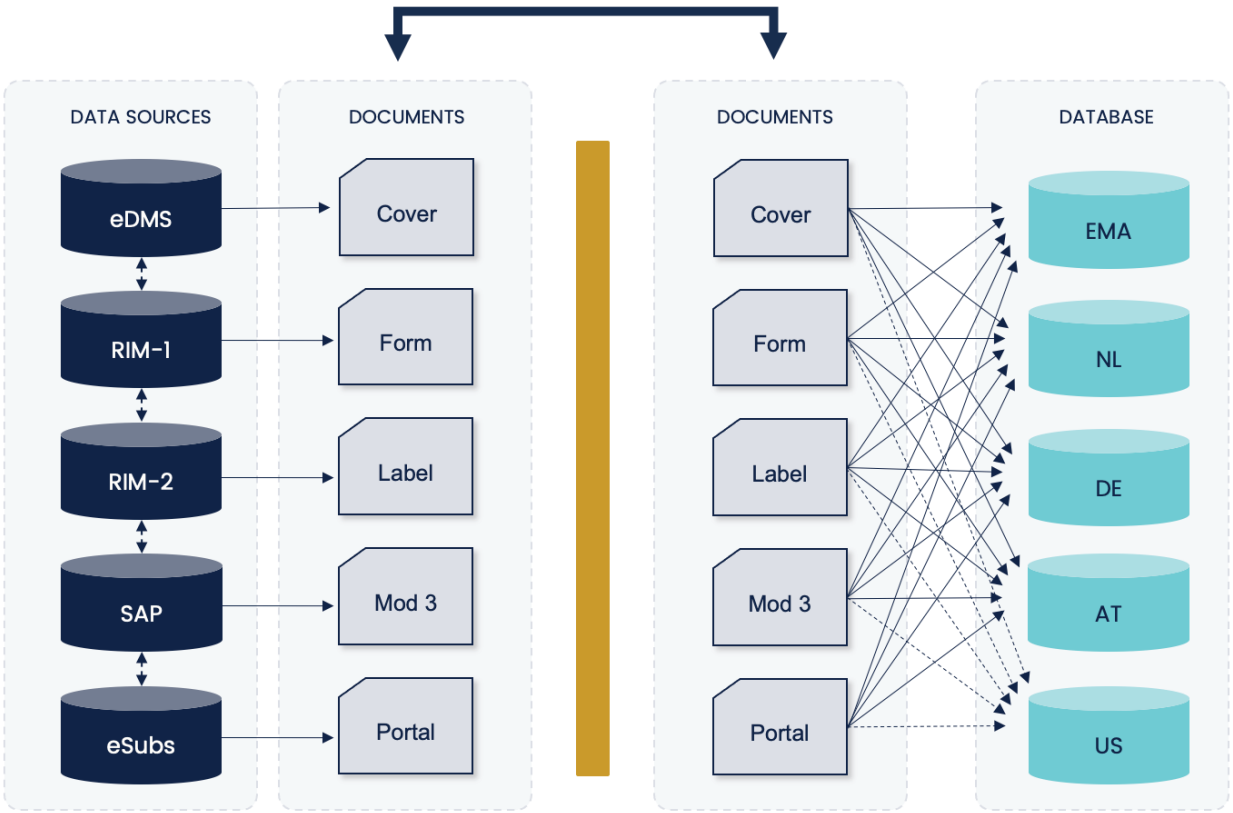
Text / Prose	Tabular Data
▪ Contextual richness, narrative depth ▪ Used to convey complex ideas, explain relationships, and provide qualitative insights ▪ Essential in communicating nuanced information ▪ Use cases: clinical findings, regulatory implications, and patient narratives	▪ Presents qualitative details well ▪ Concise and organized format ▪ Ideal for displaying numerical or categorical information ▪ Use cases: quick comparisons, trend analysis, and data aggregation; scientific research and statistical analysis

4.2. Principle 2: Focus On Content & Context

The second principle is to create the context for each data set. The data will be used as 'elements' or modules in different end documents for various HAs or therapeutics. No branding is needed in the documents other than the cover letter, application form, and product information, which are regional documents as intended.

Traditional Method – Document-Centric:

Current data, document and dossier exchange within and across industry and agencies



Source: "From CMC document to CMC data 10/10/17; Charles Morgan, Rodrigo Palacios, Hans van Bruggen

Example: Narrative vs Table Concentrate for Solution for Injection

Narrative

32P1 Concentrate for Solution for Injection

The drug product, ProQuit (Number-456) is a colourless concentrate for solution for injection containing 5 mg/ml Qdrug.

The concentrate for solution for injection is packed in a glass vial, with a minimal extractable volume of 2 ml.

The concentrate for solution for injection is to be diluted with the solvent water for injection prior to administration.

One glass vial of the concentrate for solution for injection is co-packed with one glass vial of solvent in a carton box.

One pvc/pvc blister is packed in a carton box.

Tabular View

32P1 Concentrate for Solution for Injection	
Product name:	ProQuit
Manufactured dose form:	Concentrate for solution for injection
Concentration:	5 mg/ml
Administrable dose form:	Solution for injection
Strength/Concentration:	500 mcg/ml
Active substance:	Qdrug
Colour:	Colourless (clear)
ID number:	Number-456
Immediate container:	Glass vial
Glass vial volume:	3 ml
Minimal extractable volume:	2 ml
Intermediate container(s):	pvc/pvc blister
pvc/pvc blister volume:	1
Outer container:	Carton box
Carton box volume:	1

Reference: https://gsrs.ncats.nih.gov/assets/files/GSRS_UPC_2018/Schmuff-USP%20GINAS%20GSRS%202018-11-16.pdfsee slide 18

4.3. Principle 3: Use Templates

The third principle is to apply templates. Based on years of experience with submissions, Celegence provides pharmaceutical companies with templates.

Here are some best practices related to template use:

- Keep the end game in mind and compliance-ready along the way
- Provide information only once and at the dedicated location
- Apply lean authoring, including lean cross-referencing principles
- Follow the lean authoring principles covered in Table 2, including lean cross-referencing principles, hyperlinking of cross-references to avoid obsolete links and following future changes in destination documents. Module 3 is very structured and does not need many cross-references
- Include requirements early in the document creation process

4.4. Principle 4: A Framework for Repeatable and Sustained Change

Principle 4 sets the organization up for success through maintenance:

- Establish a strong data governance that covers controlled vocabularies/terminology
- Map tables to align to other controlled vocabularies for the same concept
- Reference authoritative sources of data
- Define roles that foster cross-functional collaboration and communication
- Establish organizational accountability and support
- Training and guidelines for authors
- Invest in tools and technologies for lean authoring, such as AI-driven solutions for error detection and content management, plus dedicated dossier management

4.5. Lean Authoring Example

But what is possible now when the exchange standard is still PDF? In a PoC at one of our customers, we developed a transition stepping stone towards data models, semantics and reliability while the exchange mechanism was still in PDF.

We applied the following rules:

- Move away from narrative texts for factual data and information; provide tabular formats.
- Tables are preceded by a table title to inform the reader on the subject
- Tables consist of a table header, left-hand column and right-hand column:
  - » The table header informs the reader on what the attribute names are and its corresponding values
  - » The left-hand column lists the attribute names (or predicate)
  - » The right-hand column lists the corresponding values and subject

Example Of Table Titles, Headers, Attributes And Values

Property	Value
Company preferred name	<Company preferred name>
Manufactured dose form	<Manufactured dose form>
Administrable dose form	<Administrable dose form>
Active substance	<Active substance>
Active substance strength Choose an item. single value or low limit numerator	<Quantity operator> <Active substance strength presentation single value or low limit numerator > <Unit of measure>
Active substance strength Choose an item. single value or upper limit denominator	<Quantity operator> <Active substance strength presentation single value or upper limit numerator > <Unit of measure>
Reference substance	<Reference substance>
Reference substance strength Choose an item. single value or low limit numerator	<Quantity operator> <Active substance strength presentation single value or low limit numerator > <Unit of measure>
Reference substance strength presentation single value or upper limit denominator	<Quantity operator> <Active substance strength presentation single value or low limit numerator > <Unit of measure>
Fill volume	<Fill volume>
Minimal extractable volume	<Minimal extractable volume>
Overage	<Overage>
Shape	<Shape>
Color	<Color>
Imprint	<Imprint>
Image of the tablet	<Image>
Dimensions of the tablet	
Height	<Height> <Unit of measure>
Width	<Width> <Unit of measure>
Depth	<Depth> <Unit of measure>
Weight	<Weight>
External diameter	<External diameter> <Unit of measure>
Capsule size	<Capsule size> <Unit of measure>
Capsule shell material	<Capsule shell material>

Source: Hans van Bruggen

Interpretation of the Information within a Table

Based on the above representation of information and the logic of interpreting the information, the data is available for extraction into a database. As a result, PDF documents have become an industry standard for data exchange. This is not optimal, though; the values in MS Word can be tagged with the attribute names, rendered to PDF version 1.8, and those tags will remain in the PDF rendition. Subsequently, the information can be extracted from the PDFs. A positive Proof of Concept study was done by the IRISS Forum in 2014. At that moment, a data model was considered necessary to continue the development of that solution. Even now, there is still no data model because of the lack of agreement on a model. It is recommended that stepping stones towards a new model be accepted.

Often, the authors of regulatory documents provide too much information about a particular topic rather than convey a message. In addition, to exchange regulatory information, we still rely on the exchange of PDF files in a regulatory dossier.

4.6. Overcoming Common Challenges in Lean Processes for Pharmaceutical Submissions

Potential Challenge	Solution
Ensuring Compliance with Diverse Regulatory Requirements	Develop a comprehensive regulatory intelligence system that tracks and updates regulatory requirements for different regions. Use this system to create dynamic templates that apply to various regulations.
Managing Cross-Functional Collaboration	Use collaborative platforms that allow real-time editing and feedback. Schedule regular cross-functional meetings to align teams on project goals and timelines.
Reducing Time and Increasing Efficiency	Leverage automation tools for repetitive tasks such as format (e.g., proper use of upper case), compliance checks (whether no empty fields are present), and eCTD readiness (for allowed fonts, page layouts and cross references). Implement parallel review processes where different sections of the document can be reviewed simultaneously.
Maintaining Document Quality and Consistency	Develop a rigorous quality control process that includes a single round for the content review and a single round for the eCTD readiness check. Use standardized templates and apply text editing features to ensure consistency across documents; a text editor (like MS-Word) is more than a type writing machine.
Integrating Advanced Technology	Invest in a robust electronic content management system (eCMS) and document automation tools. Ensure seamless integration with other systems such as regulatory or enterprise information management systems (EIMS/RIMS).
Managing Change and Adoption of New Practices	Implement change management practices, including clear communication of the benefits of lean authoring, the benefits of proper templates and text editing features, and involve key stakeholders in the transition process.

By adopting these solutions, pharmaceutical companies can overcome common challenges in lean authoring, resulting in efficient, high-quality, and compliant regulatory submissions that benefit patient safety and healthcare outcomes.

> *"Improving the reliability of the data quality increases the efficiency not only within a pharmaceutical organization but also increases the efficiency from Health Authorities to a pharmaceutical organization."*
>
> **Hans van Bruggen, 2024**

5. Guide – Checklist On How To Apply Lean Authoring

Table 2. Comparison Of Traditions & Lean Authoring Principles, Including Checklist

Category	Traditional authoring Leading To CTD Documents	Lean Authoring Leading To eCTD-Ready Documents	✔
Strategy and Approach			
Communication & Collaboration	. Hand-off between teams, siloed work . Inefficient communication	. Cross-functional collaboration from the start, shared vision . Efficient communication and sharing of standardized templates	☐
Planning & Strategy	Extensive planning and outlining, often with detailed project plans and rigid timelines.	Agile planning with flexible and iterative processes, allowing for quick adjustments based on feedback and changing requirements.	☐
Stakeholder Buy-in	Stakeholders reviewed a document that contained information in context of a named purpose (such as, IND for the US or MAA for the EU). To inform the reader about the purpose, this was often mentioned in the documents and documents had to be re-written to use it for a different purpose (such as a complete document of control of drug product, including the analytical procedures, validation and justification).	Gain buy-in from all stakeholders, including writers, reviewers, and regulatory affairs teams, who can see the document for review in context from other documents that define the purpose of use, such as a dossier containing the application form or other CTD documents that provide more context around its purpose, or, the analytical procedure and its validation when reviewing the justification of the specifications.	☐
Patient Safety	The needs of patient safety and ethics transcend regional borders. Yet, the fragmented approach was not optimized for patient safety given limited comparison ability across regions and documents. The siloed approach does not benefit patients.	Create standalone documents that are reusable across regions without any rework. If different formulations and strengths are used across regions, enable the ability to pick and choose the documents that are relevant to a particular dose form or strength. Hence, identical documents can be re-used across countries to safeguard the country's patients.	☐
Human Expertise	Knowledge centralized to a small number of experts, which creates bottlenecks and overdependence on a few people.	Identify SMEs up front; review content for technical soundness, scientific accuracy, and adherence to regulatory guidelines.	☐
Version Control System	Content vulnerable to errors and quality issues due to different versions and iterations of common documents.	Reusable content modules avoids confusion or errors.	☐
Compliance	. Post-authoring compliance checks . Risk of non-compliance due to oversight	. Compliance ensured from the start with integrated eCTD requirements . Reduced risk of non-compliance with proactive measures	☐
Technology Utilization	. Limited use of advanced tools . Manual updates and maintenance . Limited tool validation	. Utilization of modern tools and RPA for authoring and submission . Automated updates and maintenance of submission tools . Ensure any new software or templates are rigorously validated to comply with regulatory requirements.	☐

Category	Traditional Authoring Leading To CTD Documents	Lean Authoring Leading To eCTD-Ready Documents	✔
Strategy & Approach			
Training	Training was narrow and specific, did not apply to approach or best practices	Training on lean authoring and eCTD v4.0 compliance	☐
Adaptability	▪ Rigid adherence to the original plan, making it difficult to adapt to new information or changing needs ▪ Focus on creation	▪ High adaptability with a focus on responding to feedback and evolving audience needs, allowing for more dynamic content creation ▪ End game in mind, lifecycle of using data	☐
Document Specific Tactics			
Language: British English versus US English	British English was often used for the EU (and UK) submissions whereas the documents were changed to US English for the US.	Considering the Common Technical Document, one can choose to use US or British English, as long as the language is consistently used within a document. The same applies for the decimal dot versus comma. Hence, in the eCTD, it is possible that one document – is in US-English, whereas another document is in British English. Module 1, containing the regional information, such as cover letter and application form and the product information (e.g. PI and SmPC) must be in the local language and use the appropriate decimal comma or dot. That means for Module 2 – 5, all documents can be reused across countries.	☐
Managing minor spelling 'errors' in regulatory documents	Upon re-using a document, spelling mistakes/ typos might be spotted. Often a new version was created to correct it prior to use, such as a document for the next submission. This often results in two slightly different versions in the document management system.	Slightly different documents (e.g., for typos), should not be updated prior to the next submission. Updating a document means that one country received a version with the typo and another without the typo. For a computer these two files are distinctly different, whereas for a human being, these are the same. Only when the typo concerns a different meaning would the documents have to be updated (e.g. 10 mg vs 1.0 mg).	☐
Font Type	In paper documents all font types were possible, as long as the printer was able to print it.	For electronic documents, the PDF should be readable in its electronic format, such as 50 years from now. The font types of Arial, Courier and Times New Roman are considered still readable for this period. Other fonts are not necessarily available in the reader's PDF browsers. To be on the safe side, the PDF needs to embed the relevant subsets of any font type used in the PDF, other than Arial, Courier, and Times New Roman.	☐
Special Characters	Often special characters are included by the author as font type Symbol. Though the author clearly sees the immediate result, and will see this properly reflected in the author's PDF, its appearance is dependent on the browser used to read a PDF. If neither the browser recognizes the font type of a character and the font types are not embedded, a special character might be shown as a box (like □) or the regular character (like 'mg' instead of µg). The latter is of course a serious difference.	What is true for the above font types is also true for font types like Symbol and Wingdings, which are often used for special characters. Hence, it is best to stick with Arial or Times New Roman, and to use 'normal text' when including special characters.	☐

Category	Traditional Authoring Leading To CTD Documents	Lean Authoring Leading To eCTD-Ready Documents	✔
Document Specific Tactics			
Heading 1 versus Title	Headings are used to structure the content of a document. The title of a document is different from Heading 1. Too often Heading 1 is misused to show the title in the bookmarks. Since only 4 level bookmarks are recommended, misusing the first level results in the use of three levels only.	Up to the fourth level bookmarks are recommended (when applicable), for large documents such as overviews, summaries and study reports. By applying the right Heading styles, the resulting PDF will present a meaningful list of bookmarks upon rendition to the PDF. Moreover, the title of a document should be part of the recurring header of a document.	☐
Section Breaks & Headers/Footers	Section breaks in MS-Word include the page layout and margins, formatting characteristics and its header and footer information for the corresponding section. The relevance of this is often underestimated.	By reducing the use of section breaks, it reduces the risks for individual changes in the formatting and header/footer content. In the event of differences in page orientation, the header and footer can be kept the 'same as previous'. Provided that the headers have been built correctly, the headers will spread out nicely, irrespective of orientation for less manual rework.	☐
Introductory Texts in Documents	Traditionally, it would not matter if a chapter started with a narrative, followed by numbered heading.	In the digital world, a change in document granularity can happen over time (e.g. limited amount of detail in the pharmaceutical development section in the IND/ CTA stage, whereas lots of information at the NDA/ MAA stage. As a result, a single document with 5 small chapters can be used for the IND, while later, it is split into five documents, each with its own structure. What were previously the names of the chapters, have now become the subtitles of a CTD section, and should become apparent in the header. To accommodate this evolution of a topic, the document should be able to be split without any rework. Be aware that there should be no 'orphan' texts at the same level as subheadings in a chapter.	☐
Document Granularity	Since the introduction of the paper CTD, documents have been authored in accordance with the granularity of ICH M4 (remainder of the regulatory dossier) and ICH E3 (for CSRs). The eCTD specifications suggested a finer document granularity, but the trend was set to a coarse granularity.	The finer the granularity, the more standalone documents can be created that can be reused for multiple purposes, such as multiple countries, products and application types. And the XML backbone of the eCTD creates a meaningful table of contents to put the documents in the right order, ensuring the correct flow of the messaging.	☐
Cross Referencing	In scientific journals and theses, many cross references must be provided to substantiate the value of the message.	Regulatory documents must always be seen in context of a common technical document (CTD). The common structure has a home for every type of regulatory information. This common structure is known by all the stake holders, hence, the need for cross referencing is limited, because the common structure defines where what information is to be provided. Automatically updated cross references can be applied during the editing phase using the feature cross reference to a named heading, caption (for tables and figures) or endnote (for reference lists).	☐
File Path For Storing Electronic Dossiers	Long path lengths were used to be descriptive without regard for application; the naming was often truncated resulting in broken links and unusable documents or references to internal sites and drives.	. Distinguish absolute and relative links . Path length used should be less than 256 characters including .pdf	☐

Category	Traditional Authoring Leading To CTD Documents	Lean Authoring Leading To eCTD-Ready Documents	✔
Document Specific Tactics			
Hyperlinking	The early non-eCTD electronic submissions did lack a common structure. Hence, the cross referencing was more often needed. Every cross reference was to be hyperlinked. Because of the lack of dossier- and document lifecycle, none of the hyperlinks would become obsolete in the future.	The eCTD allows for document replacements in a dossier lifecycle. Hence, documents can become outdated. Hyperlinks to outdated documents cannot be redirected to the replacement files. As a result, hyperlinks pointing to an outdated document remain active and carry the risks that old data is considered current.	☐
Hyperlinking Strategy	A Regulatory Operations group to create hypertext links (hyperlinks) for all cross references within and across documents of a regulatory dossier.	▪ Hyperlinks within a document: to be created implicitly by the author using the proper features for headings, captions, endnotes, footnotes and cross references. ▪ Hyperlinks between documents: a syntax to be provided by the author of where in the common structure the reader can find the referenced information. The syntax should be 'refer to', followed by the CTD title of the destination document. As a result, every reader can find the destination document in the applied dossier, even without a hyperlink. ▪ Nevertheless, to balance the ease of hyperlinks and the risk for future obsolete links, the following principle can be applied: 1. No hyperlinks to Module 1 documents 2. Deep linking to the named destination in Module 2 documents 3. No hyperlinks to Module 3 documents 4. Shallow linking to open Module 4 and 5 documents, provided that the ToC, ToT and ToF will bring the reader to the named destination using a second click	☐
Use of Colors	Colors might be shown differently by the author than by the reader. Depending on the PDF version of the document and the capability of the browser, the colors might be seen as intended.	The eCTD defines that the PDF version to be used is at least PDF v1.4.	☐
Format of Tables	The format of a table was often not limited to any structure, size or color. Which is OK for an on-off document.	Health Authorities want to copy/paste data in their own spreadsheets and databases to run their own analyses. Therefore, they have to copy/paste information from the document. For them, it is easier to get documents from tables than from narratives. The table titles and table header have to unambiguously display the purpose of the table and the subject.	☐
Location of Captions	Table captions (titles) are above the table, whereas figure captions (titles) are below the figure. This is fine for paper articles and journals.	In the electronic world, a hyperlink to a figure gets the viewer to the caption at the top of the screen. Subsequently, the viewer needs to scroll up to see the picture. Hence, the figure caption should also be above the figure, like for the tables.	☐
Location of Tables & Figures	Manual page breaks and line breaks are often used to move a table/figure and its title and key to the next page.	Automated position of the table or figure is when the object is 'in line with text'. As a result, upon authoring, it flows with the text. Provided that the caption is labelled as 'keep with next', the caption and object move at the same time across pages. The keep with next should also be applied for the tables, so an eventual table key will also be kept with the table.	☐

Category	Traditional Authoring Leading To CTD Documents	Lean Authoring Leading To eCTD-Ready Documents	✔
Document Specific Tactics			
Units of Measure	When documents are written for a particular region, the units of measure (SI versus empirical) might be accommodated to the receiving country.	For common technical documents the SI units of measure should be used. Exception is the use of 'liter' to express volume. Hence, it is not predefined how milliliter should be abbreviated (ml or mL). To prevent unambiguity between the number 1 and a lowercase l, the uppercase L is recommended.	☐
Date	The presentation of a date is done in so many different ways.	ISO has defined a standard notation for a date, namely YYYY-MM-DD. Using this, the dates can be included in a list and such a list can be chronologically ordered by any computer application, even DOS.	☐
Capitalization	Capitalization was up to the author.	Capitalization of headings are also reflected as such in the bookmarks automatically rendered in the PDF. Hence, not only the appearance of the capitals in the MS-Word document, but also how it was originally typed in. For example, the typed text might be a combination of lower and upper case and the 'style' in Word might specify that all are to be in uppercase. As a result, the Word document shows all in uppercase, while the bookmarks are as it was typed. This often goes wrong with changing the case to 'Title case", since in general, words like 'on', 'and' and 'the' should not be capitalized. Title case might frustrate your carefully written title if it includes terms like 'eCTD' and mmHg. Hence, the easiest to apply sentence case and use the font size or bold to indicate a title or specific heading.	☐
Spacing	Double spacing was something introduced by IBM in its typewriting machines used for paper documents.	Current text editors, like MS-Word automatically apply sufficient spacing between words and between sentences. If two words have to remain together use nonbreaking spaces or nonbreaking hyphens.	☐
Citing Literature	Various ways can be used in documents.	For regulatory documents it is recommended to use the MS-Word feature for endnotes. Using this feature, a list of references is automatically created and even the numbering is automatic. The Bibliographical information should be used in accordance with the Vancouver convention.	☐
Document Layout	The US required their paper dossiers to come in Letter format. The have binders, cupboards and copiers designed for Letter format. Similarly, the EU had everything designed for A4-formats.	For electronic submission, the paper size is not really important, because one can zoom in or out, depending on the browser and screen one has. Nevertheless, each document should use the page size consistently, so consistently A4 or consistently Letter. For mockups of labels, the page size might even deviate.	☐
Margins	At one time, the margins had to allow for punch holes and scribbling notes or stamping health authority information on top.	Margins are far less important now. For example, one has to have some margin to allow for transport of the paper for printing. Margins of 2.5 cm (0.98") on all sides has proven to be acceptable. This applies for portrait and landscape orientation.	☐

Category	Traditional Authoring Leading To CTD Documents	Lean Authoring Leading To eCTD-Ready Documents	✔
Document Specific Tactics			
Page Numbering	At one time, some of the documents started with Roman numbering for the first couple of summary pages, followed by the Arabic numbering. Sometimes the title pages were not numbered, and Page 1 was the page following the title page.	In the electronic world, the page number is provided by the browser. Actually, a page number is therefore no longer needed. However, if page numbers are used, they should tie to the dynamic structure of the document for easy navigation.	☐
Document Format	Long prose documents, with headers and footers mostly identical across all submission documents, containing contextual information like the brand name, complete eCTD section number, Company name and sometimes even application number	Tabular formats for factual information and leave the scientific evaluation only in documents assigned for that, such as non-clinical and clinical overview, Investigator's Brochures, Pharmaceutical and Manufacturing development sections, method validation and process validation and nonclinical and clinical study report concluding remarks sections. Note that even the latter should not cross refer to anything outside of this study. By the time the study report will be exchanged with the HAs, the scientific overviews will put the report in context of the then current thinking.	☐
	Scanned PDF documents.	RPA (Robotic Process Automation) + tabular formats to move data.	
	Long prose documents, trying to introduce as much knowledge in a single document. Academic thinking to repeat in different wording, because that adds to the strength of the story. Moreover, since the document is believed to be read in its own (also outside of the eCTD), a lot of context is given with information that has a home in other eCTD documents.	Short factual information, typical for the expected information in a specific eCTD document/section only. The context is to be deduced from the collection of documents provided in a dossier. Omit long sentences and apply the Flesch Reading Ease and/or Flesch-Kincaid Grade Level. The latter is not to be done frequently, but it reminds authors to the difficulty of understanding the narratives.	
	Text based: Large blocks of text and complex data presentations.	Visualization: charts, graphs, and tables to present data consistently.	
Data Inclusion By The Author	Manual copy/paste from data sets, summary reports and study reports into the regulatory documents. If the statistical analyses obtained from development departments can produce tables to be copied as a whole, this is preferred, but this is not always the case.	Future automated population from a database to populate tabular formats. Or even better, tabular formats the database to generate the according to predefined criteria.	☐
Data Extraction By HAs	Manual data extraction from narratives and various tables, which both partly describe the same information (e.g. test group, treatment, etc.). Furthermore, because of the repetition in conventional and academic authoring, inconsistencies might occur.	Mid-term future: automated data extraction from datasets. However, since datasets cannot yet be submitted (except for some nonclinical studies and clinical studies in the US), PDF version 1.8 or higher documents can store tags belonging to the data, that can be extracted from the PDF. Tabular formats will ease in the PDF documents (current exchange standard) the copy and processing.	
Readability	Storytelling, to introduce repetitive information out of courtesy to the reader, not appreciating that the reader has the whole dossier to read, so more than that particular document only.	Providing factual information in tabular formats with factual information. Separate scientific evaluations are to be documented in separate documents in a regulatory dossier such as Non-clinical Overview, Clinical Overview, Pharmaceutical Development, Process validation and Method validation and concluding sections in the nonclinical and clinical study reports.	☐

Category	Traditional Authoring Leading To CTD Documents	Lean Authoring Leading To eCTD-Ready Documents	✔
Document Specific Tactics			
Templates	Templates with copy-pasting the text from ICH M4 (originating from 1998), or ICH E3 (originating from 1996) to be used in electronic CTDs mandated since mid 2000's, to describe the various topics. This resulted in detailed introductions, summaries and explanations in long narratives, often summarizing what is also presented in tables and graphs.	Templates for tabular unambiguous presentation of factual information, thereby ensure the requested information will be described (provided) only once and once only, without other verbiage.	☐
Context	Long introductions in each document to provide context to the messages.	Contextual information (the purpose of sharing this document) to be derived from the cover letter and application form of a regulatory application, accompanying the various CTD documents.	☐
Formatting	Narratives of any kind of structure, suggested by a high-level guidance text from 1998.	Tabular formats suggesting a long list with potential characteristics to be provided, in line with the current information on PQ/CMC and IDMP, sometimes even beyond that.	☐
Terminology	Uncontrolled terminology: free text attribute names and attribute values, captured in an unstructured narrative.	Controlled terminology: direct, straightforward, industry standard, and use of controlled vocabularies.	☐
Structure	Linear structure: information presented in a sequential format.	Modular structure: use of headings, sub-headings, bullets, tables.	☐
Search	Difficult to search and analyze.	Simplified search, analysis, and cross-reference.	☐
Template Usage	. Custom creation . Inconsistent document formats	. Use of standardized templates tailored to regulatory requirements . Consistent document formats across submissions	☐
Post-submission, ongoing monitoring and maintenance			
Quality Control	Emphasis on thorough proofreading and multiple quality checks.	Prioritizing essential quality checks and leveraging real-time feedback to make continuous improvements post-publication.	☐
Submission Process	. Time-consuming manual preparation . Multiple versions with inconsistent formats	. Automated tools and software for submission . Standardized templates for consistency	
Regulatory Review	. Longer review times due to unstructured data . Difficulties in data validation	. Faster review times with structured, clear data . Easier validation with tabular data and clear formats	☐
Review and Feedback	Lengthy review cycles with multiple rounds of feedback, often leading to delays.	Continuous feedback loops with regular check-ins and audits. Quick iterations based on input from stakeholders and audience.	☐

6. Conclusion

In summary, embracing lean authoring practices offers the pharmaceutical industry a transformative pathway toward enhanced efficiency, compliance, and patient-centricity. By optimizing document creation, streamlining regulatory processes, and fostering collaboration, pharmaceutical companies can accelerate time-to-market for innovative therapies and ensure robust adherence to global regulatory standards. This approach improves operational effectiveness and reduces costs while demonstrating continued commitment to patient safety and healthcare quality, ultimately driving advancements that benefit industry stakeholders and patients worldwide.

The impact of AI on lean authoring, particularly within the ICH M4Q framework, is transformative, acting as a crucial stepping stone from traditional methods to a more advanced, data-driven approach. While scientific evaluations currently rely on narrative forms with limited AI involvement, the incorporation of factual data in tabular formats or databases is revolutionizing the process. This structured data can be efficiently pulled together using AI, enhancing accuracy and consistency.

As the industry moves toward AI and automation, it's important to address the "simple" fixes when it comes to lean authoring. Remember, the "target" data representation is easier to achieve when it is uniform and tabular. As AI capabilities evolve, these initial steps of integrating AI for data collation and analysis lay the groundwork for a future where AI can fully automate and optimize the authoring process, ensuring regulatory documents are both comprehensive and compliant, ultimately speeding up the approval of new drugs.

Helping You With Your Pharmaceutical Regulatory Challenges

Celegence has a wealth of knowledge to help you navigate through the most complex challenges that pharmaceutical regulations can pose.

We can assist you throughout the entire process to ensure that you and your business are compliant

For more information, reach out to us at

info@celegence.com

or contact us online at

celegence.com

celegence