



content rules®
Because content matters™

WHITE PAPER

Knowledge First:

A new mental model unlocks digitalization for pharma organizations



Table of Contents

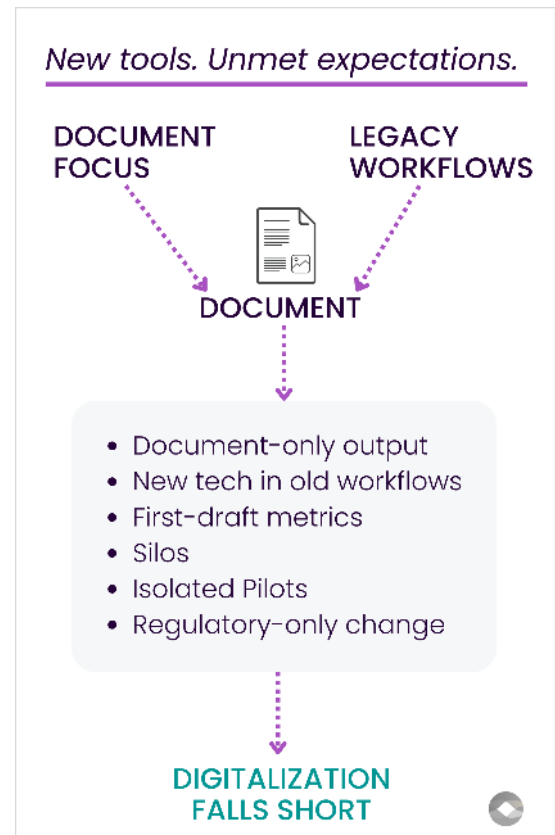
Knowledge First is the missing link.....	3
What is a mental model?.....	4
What is the Knowledge First mental model?.....	5
Knowledge is the business asset.....	6
Capture, curate, and communicate	7
1. Capture knowledge.....	7
2. Curate knowledge.....	8
3. Communicate knowledge.....	9
Locking knowledge in documents is bad for business and for patients.....	10
Risks of document-based systems.....	10
Documents limit AI reliability	12
Human-in-the-loop needs a digital approach.....	13
Digitalization requires a new mental model	15
About the authors	16
References	17
About Content Rules	18

Knowledge First is the missing link

Pharma organizations have tried several ways to leverage technology to help accelerate regulatory documentation workflows. Unfortunately, few organizations have realized the benefits they expected in shorter timelines or reduced costs.

Here are some common reasons why their expectations are not met:

- Focus on generating documents instead of on building knowledge assets
- Replicate legacy workflows in new technology
- Measure “time to first draft” for a document type instead of assessing the end-to-end knowledge lifecycle
- Allow legacy silos to block technical and data enhancements
- Operate in pilots without addressing underlying issues that prevent scalable processes
- Treat the project as part of Regulatory Affairs rather than as an organization-wide transformation



The common thread that ties these disappointed expectations together is not the technology. It's not even the process.

The number one reason that transformation fails is that there has been no lasting change to the mental model of the regulatory documentation process.

What is a mental model?

The exact definition of *mental model* varies depending upon which discipline you approach it from. In this paper, *mental model* is defined as “an interrelated set of beliefs that shapes a person's expectations for the future and how they understand the way the world works.”^[1]

Without a new mental model of how knowledge is captured, curated, and communicated, most people simply cannot make digital transformation successful. They continue to think in documents. They hold on to legacy processes instead of embracing more effective ways for their expertise to contribute to the business.

When that happens, the organization cannot realize the full potential of technology to shorten timelines and reduce risk. The return on investment (ROI) looks weak and the technology takes the blame. The digital transformation falls short of its potential or even fails altogether.

This paper describes the mental model that makes the difference between successful digital transformations that deliver business value and other attempts that do not fully pay off.

We call this mental model **Knowledge First**.

What is the Knowledge First mental model?

Knowledge First is a visionary approach to capturing, optimizing, and delivering knowledge for human use and machine processing. The Knowledge First mental model changes the picture from “we must produce documents” to “we must capture knowledge.” This approach enables large organizations to achieve business objectives more efficiently and with less risk.

In the Knowledge First mindset, writers and subject-matter experts (SMEs) focus on creating *knowledge assets* that support multiple functions within the business. A knowledge asset is a unit of information that is augmented for machine processing so that automation can be implemented to speed up processes and to reduce errors.

Knowledge assets typically comprise data, evidence, and meaning that enable the knowledge lifecycle of a drug from ideation through retirement. In addition to its role in regulatory documents and submissions, this knowledge becomes available to people throughout the organization to make decisions, evaluate outcomes, and share results.

When that knowledge is locked away in a document that is stored with hundreds (or thousands) of other documents or saved locally on a laptop, it becomes difficult or impossible to find and use that knowledge. As a result, people can make decisions and take actions based on incomplete or inaccurate information. These decisions incur risk, cost, and time that could have been avoided if the knowledge had been more readily available.

This risk increases when we bring AI into the equation.

Just about everyone knows that the quality of AI results depends almost entirely on the quality of the source information. But not everyone realizes that documents are not the optimal source format for AI to access knowledge.

Knowledge is the business asset

The crux of the Knowledge First mental model is that knowledge is the primary business asset. The document is not.

The Knowledge First mental model helps people understand that their knowledge is much more valuable to the business when it is optimized for systems to find, retrieve, and use. Their effort is rewarded when the knowledge assets they contribute are referenced by parallel and downstream systems to accelerate or de-risk different aspects of the drug lifecycle.



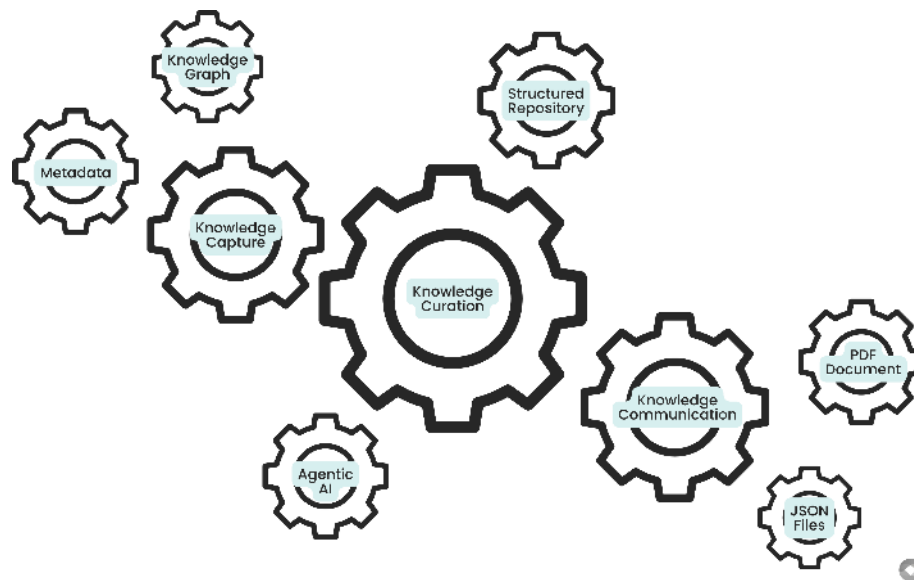
The concept of knowledge assets rather than documents is becoming more and more critical as organizations deploy agentic AI workflows. These agents are intended to take actions based on the knowledge that has been captured. Data alone is not enough to enable trustworthy agentic AI workflows; structured knowledge is key to building autonomous agentic AI at scale.^[2]

McKinsey found that only seven percent of business leaders surveyed in 2025 reported that AI was fully scaled within their organizations.^[3] The report says that scaling AI requires more than adopting new technology. According to McKinsey analysts, businesses need to redesign workflows around AI capabilities and build platforms that allow them to operate at scale.

The enterprise goal of leveraging AI to interpret information and take actions may seem removed from the daily work of any given individual involved with knowledge development. And yet, the success of any AI implementation beyond a carefully constructed proof of concept depends on making the radical shift in how the source knowledge is identified, created, and managed over the long term.

Capture, curate, and communicate

The Knowledge First mental model divides the digitalization of regulatory documentation into three categories: capture, curate, and communicate.



1. Capture knowledge

The Knowledge First mental model envisions capturing knowledge at the genesis point of the information. The knowledge is then stored and classified to enable reuse, transformation, and referencing any time the information is needed.

That information must be findable and usable for a very long time. Pharma organizations collect new knowledge starting with the first idea of a drug and

may not stop until 30 (or 50, or 100) years later. The knowledge assets that must be captured include evidence and analysis required by regulators, proprietary intellectual property about the portfolio, and all other innovations and changes that occur across the business.

Once identified, these knowledge assets need to be classified and labeled so they can be identified, discovered, and maintained. Documents simply do not allow pharma teams to label knowledge assets at the level of granularity required to enable fast, accurate search results or support AI to produce reliable results.

In the Knowledge First model, this capture and identification of the knowledge asset is a critical factor for success.

2. Curate knowledge

The Knowledge First mental model envisions knowledge curation as the ongoing process of keeping digital knowledge assets current, correct, and connected throughout the lifecycle of the drug.

It is (relatively) easy to curate the knowledge one time. The real challenge is in the ongoing maintenance of that knowledge.

Most digital transformation initiatives have focused first on automating the creation of new documents. The core project team is tasked with conducting proof of concepts or pilot programs for a specific objective (e.g., deliver the clinical data for a specific indication for a specific drug). Maintaining and updating knowledge is often a secondary concern, if considered at all.

When the use case is to update or expand the existing knowledge within the digital process, the question of knowledge curation can no longer be put off. That's when the transformation program is often abandoned as the full end-to-end knowledge curation process has not been addressed.

The Knowledge First mental model emphasizes curation as an equal priority to capturing and communicating knowledge. Curation is the hub of the entire mental model.

3. Communicate knowledge

The Knowledge First mental model envisions knowledge communication as an end result of capturing and curating knowledge.

The same digital knowledge assets can be assembled and transformed into any number of communication formats.

A document is one traditional format for communication knowledge. Other formats include XML files for clinical trial registries^[4], CDISC ODM-XML data files for electronic data capture (EDC) systems at trial sites^[5], HTML for electronic product information (ePI) websites, and structured product labeling (SPL) files for FDA submissions and the DailyMed database.^[6]

These knowledge communication formats happen downstream from document creation. Many teams are not aware of how their knowledge is used or that their knowledge contributes to these processes at all.

Today, as AI becomes more prevalent in pharma and regulator workflows, the availability of knowledge moves to the forefront of almost every business endeavor. Business and regulatory processes that used to happen “downstream” now happen in parallel or even move upstream, occurring before a document is produced. This shift in workflow goes smoother when people make a shift in how they envision the end-to-end knowledge lifecycle.

The Knowledge First mental model ensures that the people who capture or curate knowledge do not need to know all the formats that will be used to communicate knowledge. Rather, the mental model requires people to think of the communication format as an “output” of knowledge that is separate from the “input” of knowledge. We capture and curate knowledge through the inputs; we communicate through the outputs.

This separation of inputs and outputs is a relatively new concept to most pharma organizations. Yet it is one of the key distinctions between legacy workflows (which are based on how paper works) and digital workflows (which are based on how computers work).

Locking knowledge in documents is bad for business and for patients

Knowledge is the most critical enterprise asset for any pharma organization. This goes for established global enterprises with more than 100 drugs on the market and for nimble biotech startups with just one treatment in development.



Knowledge needs to be findable in order to be valuable. And yet pharma organizations regularly lock their knowledge away in the least accessible format for both humans and machines: the document.

Risks of document-based systems

Document-based systems have long been impediments to communicating knowledge. This fact is not new. It's why other regulated industries have already embraced (or at least, begrudgingly adopted) a digital approach to capturing, curating, and communicating knowledge.

Those industries now have decades of experience capturing knowledge in digital assets that can be searched, assembled, and delivered in any number of communication formats.

Dell, Inc.'s technical communications team had already been creating structured knowledge assets for more than 10 years when agentic AI became a business reality.^[7] Those knowledge assets were captured, stored, classified, and curated to enable about 60% content reuse across millions of HTML and PDF pages. In 2026, these assets became the foundation of the company's "Next Best Action" AI agent. This solution greatly speeds up customer support resolution and is credited **with an estimated \$240 million in support cost deflection.**^[7] Their workflows reflect the Knowledge First mental model and quickly adapt to support AI solutions.

As an industry, pharma organizations have struggled to effect a similar radical change in how knowledge is captured, managed, and delivered.

As a result, pharma organizations cannot:

- Amass knowledge into an accessible place while also preventing redundancy, contradictions, and inconsistencies
- Keep information current and compliant without extensive manual work that is often hidden and not documented
- Leverage automation to provide the right content in the right format
- Provide enough context for AI to find the exact right information in a document library riddled with redundancy and potential contradictions

Unfortunately, some of those organizations have attempted to address these risks by implementing AI without taking a critical look at the knowledge foundation.

Documents limit AI reliability

We can all see the potential of AI and content automation to help get medicines to patients faster. And, we have seen over and over again that new tools alone do not bring as much acceleration to content development as desired.

AI cannot add value when the knowledge is locked away as unstructured documents.

Every large pharma organization we speak to is looking to AI as a way to speed up document-based processes. However, the risks inherent in document-based processes increase, not decrease, when you add AI without also changing how you capture and curate knowledge. In fact, according to Gartner, more than 40% of agentic AI projects are likely to be canceled by the end of 2027.^[8]

Nobody knows just how many errors, gaps, and outdated information exist in the vast document repositories that represent the accumulated knowledge of a drug. That was accepted as an inevitable part of the document-centric process when the risks lay in the chance that (1) a person encountered one of these issues and (2) it caused enough of a problem to require escalation.

When AI is incorporated into document-based processes, the errors, gaps, and outdated information within the thousands of documents become obvious on a scale that wasn't conceivable with search results and human assessment. These issues impact major areas such as drug ideation, safety profiles, and analysis of related drug substances and products.



In test after test, **structured knowledge assets** enriched with context and explicitly defined relationships **outperform static documents** in any AI-enabled processing.

30%–60% Faster

Structured graph queries provide 30% to 60% faster initial generation than ad-hoc chunk retrieval.^[9]

Context Matters

RAG retrieval suffers when source content lacks context, metadata, and explicit relationships.^[10]

Structured Data Required

It is impossible to automate classification tasks reliably without structured data.^[12]

Less Training Data

AI models trained on well-labeled and semantically defined content perform better with less training data.^[11]

With AI, there is no hiding those deficiencies. Can you detect them before they propagate and escape to the outside world? What are the repercussions to patient safety? To enterprise reputation? To the bottom line?

Every pharma organization expects to use AI to transform source knowledge assets into additional assets for use in submissions, annual reporting, patient-facing materials, and other outputs. The expectation is that AI will help pharma teams meet knowledge communication requirements and obligations on a global scale.

This expectation can only be realized when AI applications and agents have access to curated, structured knowledge assets.

Human-in-the-loop needs a digital approach

One common approach to digitalization is to incorporate “human-in-the-loop” (HITL) review as the safety net for AI-generated documents.

The total reliance on HITL is an industry-wide admission that we cannot trust AI to do the job correctly when we use AI to generate documents from unstructured knowledge sources. However, Dell’s content engineering approach suggests that when humans focus on capturing and curating structured knowledge assets, then AI can produce much better results.^[7]

The combination of AI + HITL is supposed to cut weeks and months out of the document development process. It can, but not if you just use documents to create more documents. The process of reviewing information still requires review of source documentation to ensure that the analysis and conclusions are consistent and complete. People must also perform close review of the language and details so that the submission presents a clear, accurate, and persistent view of the data and conclusions.

A digital approach that puts structured knowledge assets at the center is the only way to achieve process acceleration and content quality at scale.

Organizations that have achieved significant acceleration in the production of clinical study reports (CSR) and CMC documentation did not achieve these results simply from adding AI into document production. Rather, they all changed their fundamental approach from documents to structured knowledge assets. They adopted new practices such as lean authoring, structured content authoring, structured content and data management, and lean templates.

KNOWLEDGE FIRST IN PRACTICE

Organizations didn't simply add AI.
They changed the way knowledge was created and managed.

 **NOVARTIS**

25%

Reduction in time to registrational CSR, with an average writing effort of 5.25 weeks.^[11]

 **Pfizer**

50%


novo nordisk®

90%

 **MERCK**

50%

Reduction in CSR document length

Removes thousands of words for AI to parse and humans to review^[12]

AMGEN

Structured Knowledge

Automation of regulatory authoring

Uses semantically structured blocks with metadata defining relationships between scientific data and business contexts^[13]

Digitalization requires a new mental model

Adding new technologies into existing document-first processes or undertaking single-purpose pilots that do not scale will not equate to successful digitalization. To realize the full ROI requires a significant change in how people work. For people to accept these changes, they must develop a new understanding of digitalization.

To be successful at scale, pharma must fundamentally change its current approach to digitalization. It is evident that the current approaches are not providing the desired results.

The fundamental change required is to change the way we think about knowledge itself. Knowledge First is the mental model that makes that shift possible.

The Knowledge First mental model helps the people closest to the business-critical knowledge develop a common vision. This unifying vision then enables them to adopt new technologies, processes, and workflows that empower successful digitalization.



About the authors



Kathleen Vieson

Kathleen Vieson, PharmD, CMPP, is a life sciences content strategy and digital transformation consultant. She helps pharmaceutical companies and content technology vendors transform clinical and drug information through structured content, governance, technology implementation, and AI readiness. A clinical pharmacist and oncology specialist by training, Kathleen brings deep expertise across healthcare information, content operations, and technology. She has led large cross-functional content teams, developed structured content and reuse strategies, and built content and data management systems that support scalable digital delivery. Today, she advises global pharma organizations on moving from document-centric processes to structured, reusable knowledge ecosystems.



Regina Lynn Preciado

Regina Lynn Preciado is the VP of Content Strategy Solutions at Content Rules, Inc. She leads content strategy teams to help pharma organizations optimize content workflows and maximize the business value of their investment in structured content ecosystems. Regina has helped clinical, non-clinical, labeling, quality, medical device, and marketing teams mitigate risk and enable sustainable content practices that scale. Regina is an industry expert with more than 25 years of experience in structured content authoring, component content management, AI readiness, and content reuse and automation. She lives a dogspotting lifestyle.

References

- [1] Holtrop JS, Scherer LD, Matlock DD, Glasgow RE, Green LA. The Importance of Mental Models in Implementation Science. *Front Public Health*. 2021 Jul 6;9:680316. doi: 10.3389/fpubh.2021.680316. PMID: 34295871; PMCID: PMC8290163. <https://pmc.ncbi.nlm.nih.gov/articles/PMC8290163/#s2>
- [2] Talisman, J. (2025, June 30). Why AI isn't autonomous (yet), Part III: What agentic AI needs. Intentional Arrangement. *Substack*. <https://jessicatalisman.substack.com/p/why-ai-isnt-autonomous-yet-ea2>
- [3] McKinsey & Company. (2025, November). AI at work but not at scale. *McKinsey*. <https://www.mckinsey.com/featured-insights/week-in-charts/ai-at-work-but-not-at-scale>
- [4] National Library of Medicine. Accessed August 13, 2026. Steps for registering a clinical study: Register through XML upload. *ClinicalTrials.gov* <https://clinicaltrials.gov/submit-studies/prs-help/how-register-study>
- [5] Sam Hume, CDISC, State College, PA, USA; Jeff Abolafia, Rho, Chapel Hill, NC, USA; vGeoff Low, Medidata Solutions, London, UK; Use of HL7 FHIR as e Source to Pre-populate CDASH Case Report Forms using a CDISC ODM API. *CDISC*. <https://www.cdisc.org/kb/articles/use-hl7-fhir-esource-pre-populate-cdash-case-report-forms-using-cdisc-odm-api>
- [6] National Library of Medicine. Accessed August 13, 2026. *DailyMed*. <https://dailymed.nlm.nih.gov/dailymed/fda-drug-guidance.cfm>
- [7] Eric Brehmer (2026). The impact of agentic AI on technical documentation: How Dell is responding. Tridion Live Sessions, <https://www.rws.com/content-management/tridion/resources/tridion-live-sessions-2026/>
- [8] Gartner, Inc. (2025, June 25). Gartner predicts over 40% of agentic AI projects will be canceled by end of 2027. *Press release*. <https://www.gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent-of-agentic-ai-projects-will-be-canceled-by-end-of-2027>
- [9] Lantosca, M. (2025, October 26). Using knowledge graphs to accelerate and standardize AI-generated technical documentation. *Medium*. https://medium.com/@nc_mike/using-knowledge-graphs-to-accelerate-and-standardize-ai-generated-technical-documentation-for-03f07b8df585
- [10] Choubey, P. K., Peng, X., Bhagavath, S., Huang, K.-H., Xiong, C., & Wu, C.-S. (2025). Benchmarking deep search over heterogeneous enterprise data. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing: Industry Track* (pp. 501–517). Association for Computational Linguistics. <https://aclanthology.org/2025.emnlp-industry.34/>
- [11] Geary, L., Nielsen, I.Ø., Charron, F., Carter, J. J., Martinsson, L., Rowe, E., Goese, A., Chaturvedi, A., Gidh-Jain, M., Hreha-Phillips, S., & April, J. (2026). Optimizing clinical study report development through lean authoring: A framework for meeting regulatory expectations and preparing for automation. *Therapeutic Innovation & Regulatory Science*. <https://link.springer.com/article/10.1007/s43441-026-00980-6>
- [12] Brown E, Jochman K. Regulatory matters. *Med Writ*. 2021;30(1):69–71. <https://emwa.org/wp-content/uploads/Media/media/3762/regulatory-matters.pdf>
- [13] Ahluwalia, K., Abernathy, M. J., Beierle, J., Cauchon, N. S., Cronin, D., Gaiki, S., Lennard, A., Mady, P., McGorry, M., Sugrue-Richards, K., & Xue, G. (2021). The future of CMC regulatory submissions: Streamlining activities using structured content and data management. *Journal of Pharmaceutical Sciences*. <https://doi.org/10.1016/j.xphs.2021.09.046>

About Content Rules

Content Rules, Inc. helps pharma organizations accelerate adoption and maximize the return on their investment in digitalization. The world's largest and most innovative companies trust us to develop their enterprise content strategies, transform their content ecosystems, optimize content for a worldwide audience, and develop effective content that gets results.

Since our founding in 1994, we have refined our approach over decades of working with organizations facing challenges just like yours.

We provide:

- Knowledge First training, coaching, and mentoring
- Content strategy and AI readiness assessments
- Content architecture and reuse strategy design
- Writer and reviewer enablement
- Governance and workflow design
- Pilot and implementation support
- Long-term adoption guidance

Contact us at info@contentrules.com for a free consultation.

