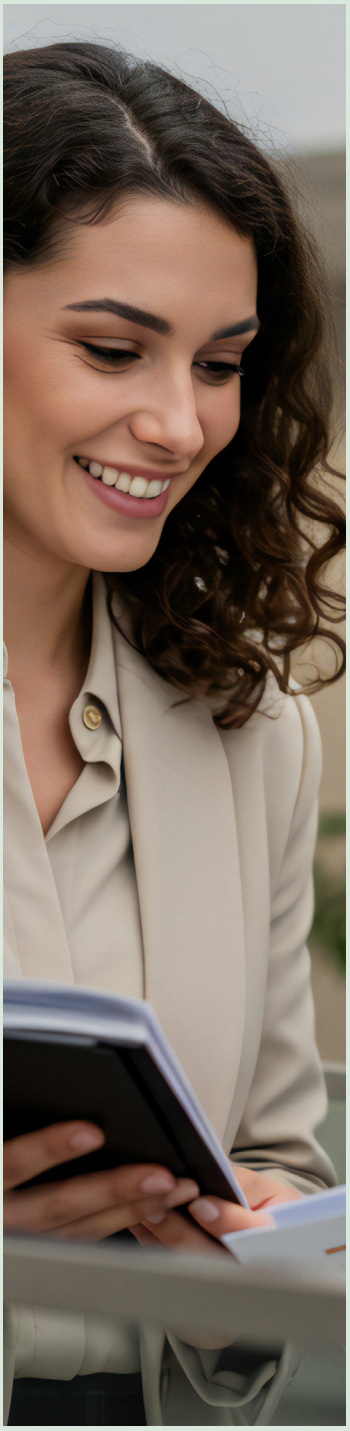
WHITE PAPER



AI-POWERED DOCUMENT PUBLISHING: ACCELERATE PUBLISHING CYCLES AND ENHANCE ACCURACY THROUGH A CONNECTED DIGITAL THREAD

Abstract

Industrial enterprises continue to rely on manual, fragmented documentation processes that remain disconnected from authoritative systems such as PLM, ERP, CAD, and MES. This white paper introduces Autonomous Digital Document Publishing (DDP) as a paradigm shift—treating documents not as manually authored artifacts, but as real-time computed outputs of enterprise data.

The DDP framework enables business-managed, low-code workflows that automate document generation, formatting, approval, and multi-format publishing at scale. By integrating structured engineering data, reusable templates, and AI-driven semantic enrichment, the platform ensures accuracy, consistency, and alignment with the product lifecycle. Governed workflows enforce traceability, compliance, and auditability, while PDF/A standards ensure long-term regulatory validity.

A validated HVAC implementation demonstrates transformative impact—reducing document turnaround from business months to Hours, eliminating IT dependency for template changes, and achieving near-zero error rates. DDP also acts as a core enabler of Digital Thread, ensuring documents continuously reflect the current state of product knowledge.

This paper positions DDP as a foundational capability for modern industry, enabling scalable, compliant, and intelligent documentation aligned with real-time engineering systems.

Infosys
Navigate your next

1. Executive Summary

Across every capital-intensive industry — manufacturing, aerospace, energy, medical devices, and utilities — enterprises maintain vast ecosystems of product-related documents. Despite the sophistication of modern engineering systems, the process of creating and maintaining these documents remains largely manual: data is exported, re-typed, formatted, emailed for approval, and published through human effort that is slow, error-prone, and impossible to scale.

Digital Document Publishing (DDP) fundamentally changes this. By treating documents as computed outputs of authoritative systems — PLM, ERP, CAD, MES — rather than as separately authored artifacts, DDP eliminates manual transcription, enforces a single source of truth, and enables one-click, multi-format document generation with governed approval workflows.

The validated outcomes for an HVAC manufacturing unit implementation establish the benchmark:

Dimension	Legacy State	DDP Platform
Turnaround Time	Business Months	Hours
IT Involvement	Every template change	Zero — business self-serves
Output Formats	Single (Word)	Multi-format (DOCX, PDF/A, HTML)
Approval Process	Email chain, no audit trail	Governed workflow, full audit log
Error Rate	High (manual re-entry)	Near-zero (PLM-sourced)
Lifecycle Currency	Periodic manual sync	Always reflects approved PLM data
Regulatory Readiness	Manual compile per audit	Automated, version-controlled

The specific technologies mentioned in this white paper for Digital Document Publishing include PLM packages → REST API data extraction → Aspose rendering engine → veraPDF validation → PDF/A-2b output, with the entire workflow governed within the PLM native approval engine.

2. Introduction

Every physical product that ships from an industrial facility carries a document ecosystem — technical specifications, compliance declarations, work instructions, operation and maintenance manuals, and regulatory submissions. These documents are not peripheral: they are regulatory requirements, contractual obligations, safety instruments, and liability shields, traceable and version-controlled across product lifecycles spanning 20 to 40 years.

Yet their production remains pre-digital. Engineers export data to spreadsheets. Technical writers reformat it in Word, and approvals travel via email. Documents are versioned with filename suffixes. Every design change triggers a manual reconciliation cycle. The consequences are well-documented: documentation backlogs that outlast design cycles, version drift between engineering data and field documents, and compliance exposure from unapproved content reaching the field.

CENTRAL THESIS

DDP treats documentation as a real-time, computed projection of the engineering system of record — not as a separately authored artifact. When data changes in PLM, the document changes. This is not incremental improvement — it is a categorical architectural shift.

3. The Industrial Document Problem

The challenges that make manual document authoring untenable at scale are consistent across industries. They share a common structural root: the document is decoupled from the data.

Version Drift Across Long Product Lifecycles

Industrial products evolve continuously through design change orders, supplier substitutions, and regulatory updates. In a manual process, each change fails to propagate immediately to all associated documentation, creating a persistent delta between the engineering record and the published document — generating field safety risk, compliance exposure, and warranty liability.

Formatting and Structural Complexity

Industrial documents contain merged-cell tables, configuration-conditional steps, and hierarchical cross-references. These requirements are manageable in a template engine but prohibitively expensive to reproduce consistently through manual authoring across thousands of document instances.

No Governed Approval Process

Email-based approval is ungoverned by design: no system-enforced routing, no SLA visibility, no escalation logic, and no tamper-evident audit trail. For products subject to regulatory oversight (FDA, CE, ATEX, AS9100), this is not merely inefficient — it is a compliance failure mode.

4. What is Digital Document Publishing (DDP)

Digital Document Publishing is the automated creation, management, versioning, and distribution of product-related documents directly from authoritative systems of record — PLM, ERP, CAD, and MES — instead of manually authoring them in Word or PDF editors. In a DDP model, a document is not authored — it is computed. The template defines structure. The system of records provides data. The rendering engine assembles the output. Human involvement is reserved for template governance and approval decisions — not data entry or transcription.

Core Platform Capabilities

Capability	Description
Data Binding	Live pull from PLM/ERP/MES — no manual transcription
Template Management	Business-owned, version-controlled — no IT involvement
Multi-format Output	DOCX, PDF/A, HTML from a single generation invocation
Approval Workflow	Governed routing with SLA enforcement and audit trail
Revision Management	Document versions linked to product revisions in PLM
Conditional Content	Config-specific sections, variant-aware tables, region blocks
Archival Compliance	PDF/A output for long-term legally valid preservation
Digital Thread	Documents connected to the product lifecycle data model

5. Why DDP is Foundational to Modern Industry

Four distinct imperatives — each independently compelling and collectively decisive — establish DDP as a platform-level investment across regulated industries.

Complex Products Across Long Lifecycles

Industrial products accumulate thousands of engineering changes across decades-long service lives. Every change must propagate to every affected document — immediately, accurately, and with traceability. A DDP platform makes document currency a system guarantee: when the PLM record is updated and approved, the document is updated — without human intervention.

Faster Time to Market

DDP enables parallel documentation: documents are maintained throughout the design process, not just at its completion. When design freezes, documentation is already 90% complete. The final generation and approval cycle is completed in hours, not weeks — a strategic differentiator for organizations competing on launch speed.

Regulatory and Compliance Requirements

A DDP platform satisfies regulatory requirements by architecture. ISO 9001, AS9100, FDA 21 CFR Part 11, ATEX, CE, REACH, and RoHS requirements are all met: version control is intrinsic; approval authority is workflow-enforced; audit trails are tamper-evident; and PDF/A output satisfies long-term archival mandates.

After-Sales Service and Customer Experience

DDP generates configuration-aware service documentation as a natural by-product of manufacturing data binding. A field service engineer requesting the manual for a specific serial number receives a document reflecting the exact configuration of that asset — with correct part numbers, parameter ranges, and procedure variants — always current, always accurate.

6. Cross-Industry Applications

The same architectural pattern — PLM/ERP data sourcing, template-governed rendering, multi-format output, governed workflow, PDF/A archival — applies across every capital-intensive regulated industry. The table below summarizes the application breadth and industry-specific drivers.

Industry	Primary Driver	Key Document Types	Governing Standards & Regulations
Industrial Manufacturing	Product complexity & variant scale	Spec sheets, work instructions, service manuals	ISO 9001
Aerospace & Defense	Airworthiness traceability	AMM, CMM, IPC, SRM, ALS	AS9100, DO-178C
Energy & Utilities	Long asset lifecycle, safety	O&M manuals, safety cases, commissioning docs	IEC 61508, ISO 55001
Automotive	Volume & variant complexity	Type approval, PPAP, owner's manuals	IATF 16949
Medical Devices	Regulatory rigour, DHF/DMR	IFU, risk files, DHF, PMS reports	ISO 13485, EU MDR
Oil, Gas & Chemicals	Operational safety	Operating procedures, SDS, MOC records	IEC 61511, ATEX, GMP

7. The Role of PDF/A in Digital Document Publishing

The same architectural pattern — PLM/ERP data sourcing, template-governed rendering, multi-format output, governed workflow, PDF/A archival — applies across every capital-intensive regulated industry. The table below summarizes the application breadth and industry-specific drivers.

DEFINITION	PDF/A is an ISO-standardized subset of PDF (ISO 19005) designed for long-term archiving and preservation of electronic documents. It ensures identical future rendering, eliminates external dependencies, and satisfies regulatory and compliance requirements.

Standard PDF is not designed with long-term archival or legal admissibility in mind. In regulated industries, documents are subject to retention requirements spanning product lifetimes — 10 to 60+ years. PDF/A provides this guarantee through five mechanisms:

- Visual fidelity — identical rendering regardless of OS, viewer version, or platform, now or in three decades.
- Self-containment — all resources embedded; no external font or asset dependencies.
- Regulatory compliance — FDA eSub, EU MDR, and national regulatory archives require or strongly recommend PDF/A conformance.
- Legal admissibility — combined with a qualified electronic signature (QES) under eIDAS, PDF/A creates an immutable, cryptographically verifiable record.
- Platform independence — no proprietary software licensing risk across a multi-decade retention horizon.

IMPLEMENTATION NOTE	PDF/A conformance must be validated, not assumed. veraPDF (open-source, ISO-endorsed) should be integrated into the generation pipeline to validate every output against ISO 19005 before archival. Common failure modes — non-embedded fonts, device-dependent color spaces, missing XMP metadata — are all detectable and correctable at generation time.

8. Digital Document Publishing and the Digital Thread

The Digital Thread is the connected flow of information linking every stage of a product's lifecycle — from initial concept through engineering design, manufacturing, quality, service, and retirement. DDP is the mechanism by which the Digital Thread's information model is materialized as human-readable documents at every lifecycle stage.

DDP is not a peripheral tool connected to the Digital Thread — it is a core node. At the design stage, engineering specifications are generated directly from CAD/PLM. At manufacturing, work instructions are generated from PLM-defined processes. At delivery, configuration-specific customer packages reflect the exact as-built configuration. At service, maintenance manuals reflect the as-maintained record. At retirement, decommissioning records are generated from the full asset history.

This architectural distinction is decisive: DDP documents are not static files stored in folders — they are dynamic projections of the Digital Thread's current state, generated on demand. The document is the output of the thread, not a parallel artifact that must be manually kept in sync with it.

9. How It Works: End-to-End Document Generation

The end-to-end process operates across five system-managed stages. Human touchpoints are limited to template governance and approval authority — not data entry, formatting, or distribution.

Stage 1 — User Initiates Generation

A business user selects the target product configuration, specifies the document type, and clicks to generate from within the PLM interface. For automated triggers (ECO approval, product revision, and release), the document is generated without any human action.

Stage 2 — Data Extraction and Assembly

The platform interrogates connected systems of record: PLM for engineering and configuration data; ERP for commercial data; CAD for dimensional attributes; MES for manufacturing parameters. The assembled data object is validated against the document schema before rendering, catching data quality issues at source.

Stage 3 — Template Resolution and Rendering

The applicable template is resolved by document type, product family, and locale. The Aspose rendering engine binds the data object to the template, handling structural complexity: merged cells, conditional content blocks, dynamic table expansion, and brand element injection. Output: a fully formatted DOCX, produced in seconds.

Stage 4 — Governed Approval Workflow

The generated document enters a governed workflow. Reviewers receive system notifications. Every action — review, approval, rejection, escalation — is timestamped, attributed, and logged in the tamper-evident audit trail. SLA rules enforce review windows and trigger automatic escalation. This replaces email approval with a system-enforced, auditable process satisfying ISO 9001, AS9100, FDA 21 CFR Part 820, and ISO 13485 requirements.

Stage 5 — Multi-Format Publication and Archival

On workflow completion, the approved DOCX is rendered to PDF/A-2b, validated against ISO 19005, and distributed to configured channels. The complete cycle — from user initiation to archived PDF/A — completes in under complete in minutes for standard documents.

10. Platform Architecture

The DDP platform is structured in four logical layers with defined responsibility boundaries. This separation enables independent evolution of the data model, template layer, rendering engine, and workflow layer — critical for a platform serving multiple industries, document types, and system integrations simultaneously.

Layer 1 — Data Sourcing

REST API or native SDK integrations with PLM platforms and major CAD platforms extract item attributes, BOM structures, dimensional data, and commercial attributes. Schema validation of the assembled data object prevents malformed documents from reaching the rendering stage.

Layer 2 — Template Management

A business-owned, IT-independent environment for template creation, modification, versioning, and publication. Low-code WYSIWYG editor with automated binding validation. Templates are organized by document type, product family, and locale, with inheritance support for variant templates.

Layer 3 — Generation Engine

The Aspose rendering core manages complex industrial formatting—merged cells, rotated headers, conditional rows, dynamic sections, and cross-references. A single invocation generates DOCX, PDF/A-2b, and HTML outputs, with mandatory veraPDF validation before archival.

Layer 4 — Workflow and Distribution

Configurable approval workflows per document type and product family. Sequential and parallel routing, SLA enforcement, and automatic escalation. Distribution orchestrator triggers downstream publication to SharePoint, Documentum, Windchill, Salesforce, and bespoke portals on final approval.

Layer	Core Components	Integration Points	Key Standard
Data Sourcing	PLM/ERP/CAD connectors, schema validator	Aras, Teamcenter, CATIA, NX	REST, OData, SOAP
Template Management	Low-code editor, template repo, binding validator	Business users (no IT)	Version control, JSON schema
Generation Engine	Aspose, multi-format renderer, PDF/A validator	PLM UI, batch triggers, event webhooks	ISO 19005, OOXML
Workflow & Distribution	Workflow engine, audit trail, distributor	Portal, DMS, CRM, regulatory archives	eIDAS, 21 CFR Pt 11
Medical Devices	Regulatory rigour, DHF/DMR	IFU, risk files, DHF, PMS reports	ISO 13485, EU MDR
Oil, Gas & Chemicals	Operational safety	Operating procedures, SDS, MOC records	IEC 61511, ATEX, GMP

11. Autonomous Digital Document Publishing Workflow Architecture view:

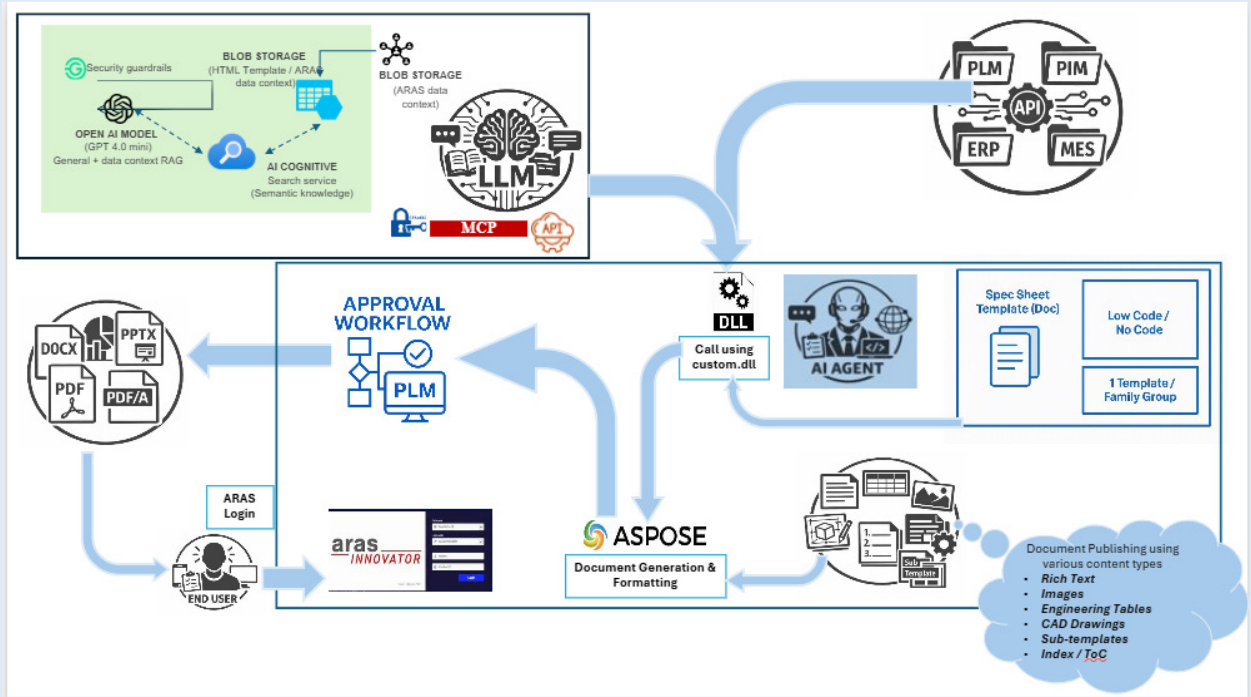
This section shows our in-house engineered capability for Autonomous Digital Document Publishing (DDP), representing the next evolution beyond conventional document automation.

The proposed architecture is a cloud-native, AI-driven workflow pipeline that enables the intelligent generation, validation, and lifecycle management of engineering documents—transforming them from static deliverables into dynamic, continuously evolving digital assets.

Built on proven Digital Document Publishing foundations and extended with advanced AI capabilities, this architecture establishes a scalable, enterprise-grade framework for autonomous document generation across the organization.

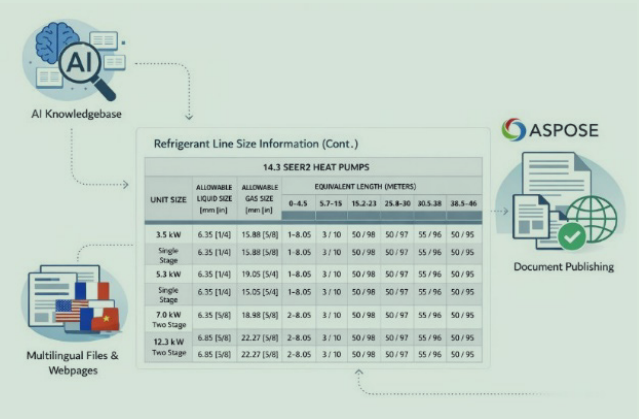
The Autonomous DDP architecture introduces:

- **Event-driven orchestration**
  Documents are triggered and updated based on real-time enterprise events (design changes, BOM updates, compliance triggers), ensuring synchronization with the live digital thread.
- **Context-aware intelligence using AI/LLMs**
  Leveraging LLMs and retrieval-based mechanisms, the system understands engineering context, enabling intelligent content construction rather than simple data-to-template mapping.
- **Closed-loop validation and continuous evolution**
  AI agents continuously validate documents against business rules, PLM workflows, and regulatory requirements, enabling self-correction and automatic regeneration when discrepancies or changes are detected.



A cloud-native, AI-driven architecture that autonomously generates, validates, and publishes engineering documents from enterprise data (PLM, ERP, MES, PIM). Data is accessed via APIs, contextualized using LLM + RAG, and orchestrated through MCP-enabled interfaces and AI agents to trigger document generation and validation workflows. Templates and configuration rules define structure, while Aspose enables dynamic composition and multi-format rendering. Governance is ensured through DLL-based business rules and PLM-driven approvals. The solution delivers structured, multi-format outputs (PDF, DOCX) with tables, images, and CAD content, transforming manual documentation into a scalable, governed, near real-time publishing pipeline.

Advantages of Engineering Digital Document Publishing



- Complex standardized table structures, which are difficult to achieve through dynamic coding
- The LLM monitors formatting so that accurate formatting, including common data merging and consolidation, is implemented without the need for code, as the LLM analyzes and formats content intelligently.
- Multilingual files and web pages can be created and attached to item types for greater dynamic clarity.
- AI search gives the LLM a knowledge base to create accurate and standardized table structures without hallucinations in the generated output.
- The Aspose package handles document publishing.

12. Case Study: Industrial DDP Implementation

The following section documents a validated DDP implementation delivered by Infosys Engineering Services for a leading American HVAC equipment manufacturer — a reference implementation that can be replicated across all industrial domains described in this paper.

Client Context

A major industrial HVAC manufacturer required spec sheet generation across multiple product families, each with dozens of SKU variants. The existing process required business months per product variant, with every template change requiring IT involvement. Systems of record: PLM Package and an external PIM system.

Solution Architecture

PLM as system of record. Low-code template management with business-owned governance. Aspose-based rendering engine with custom. dll PLM UI integration. REST API integration with PIM for commercial data merge. Multi-format output: DOCX for approval workflow, PDF/A-2b for distribution and archival. PLM-native approval workflow with SLA enforcement and tamper-evident audit trail.

Measured Outcomes

Metric	Before	After
Spec sheet turnaround	Business Months	Hours
IT involvement (templates)	Required — PLM config team	Zero — business self-serves
Output formats	Single (MS Word)	DOCX + PDF/A-2b
Approval process	Email chain, no audit trail	Governed workflow, full audit log
Data accuracy	Manual transcription errors	Near-zero — PLM-sourced
Brand consistency	Author-dependent	Template-enforced

Replication Guidance

The architecture is intentionally generic. The only industry-specific configuration elements required to replicate it in a different domain are: PLM/ERP connector mapping; product-group data representation model; domain-specific document templates (authored in the same low-code environment); workflow configuration reflecting regulatory routing requirements; and PDF/A conformance level selection (typically PDF/A-2b; PDF/A-3b for sectors requiring embedded files).

13. PLM Centric Use Cases (Engineering & Lifecycle)

Core Theme: Engineeringdriven document generation powered by PLM data backbone

Key Use Cases

Engineering Specifications / Datasheets
Auto-generate specifications directly from CAD models, BOM structures, and product attributes within PLM systems

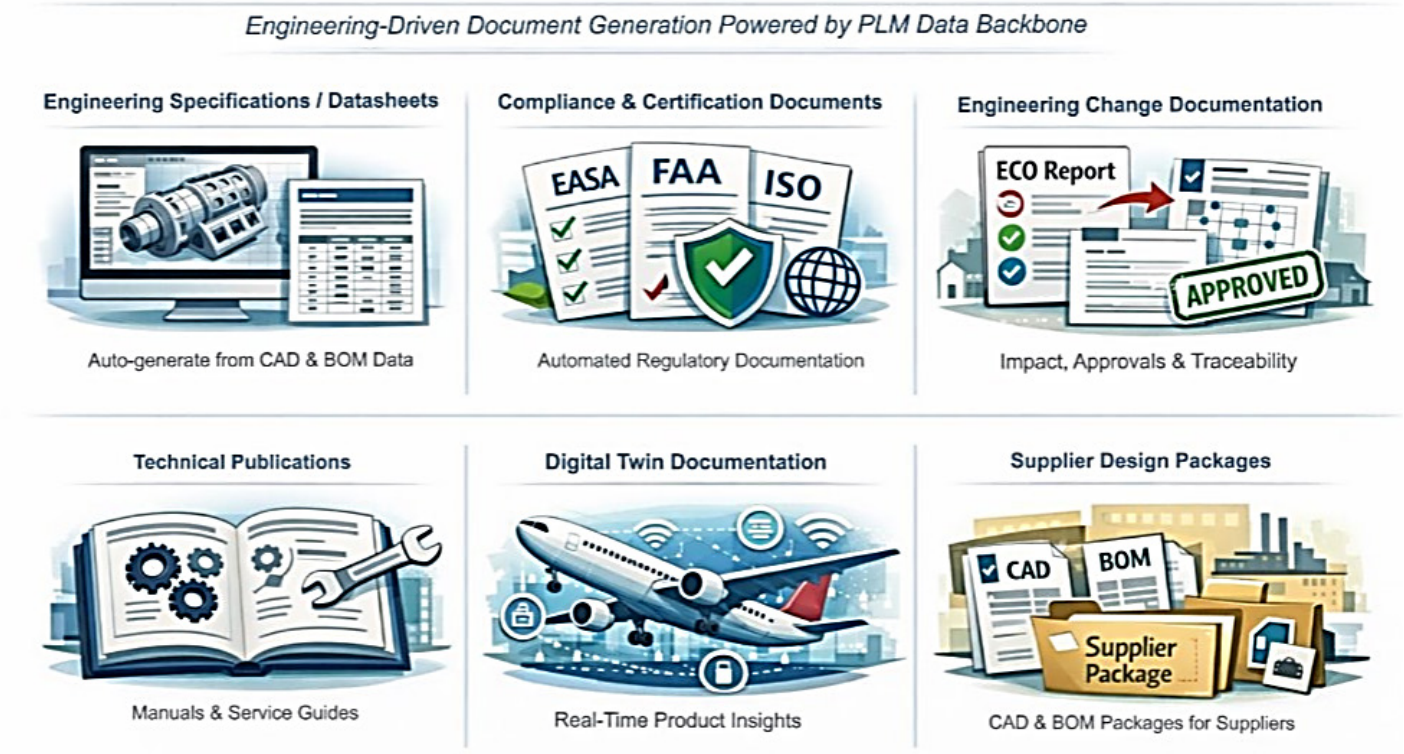
Compliance & Certification Documents
Automated generation of regulatory documentation (EASA, FAA, ISO, etc.) leveraging traceable product and audit data managed in PLM

Engineering Change Documentation
ECO / ECN reports with impact analysis, approvals, and traceability generated from PLM change workflows

Technical Publications
Creation of manuals, service guides, and assembly instructions using structured engineering data and lifecycle content

Digital Twin Documentation
Real-time documentation linked to product instances using sensor-driven digital twin data integrated with PLM

Supplier Design Packages
Automated packaging of CAD, BOM, and design documents for supplier collaboration and manufacturing handover



14. Implementation Considerations

Data Quality Prerequisites

A DDP platform is only as accurate as the data in the system of recording. Before deployment, organizations must conduct PLM/ERP data quality audits: completeness assessments, mandatory field enforcement, and revision governance review. Data quality debt is not masked by automation — it is exposed at generation time with structured error reports.

DATA QUALITY FIRST	Organizations deploying DDP without addressing PLM data quality will produce automated documents that are consistently wrong. Data quality remediation is a prerequisite, not a parallel workstream.

Template Governance

Business ownership of templates requires a formal governance model: designated template owners, mandatory validation gates, and a regular review cadence — to prevent template proliferation, version confusion, and brand or regulatory drift.

Phased Rollout Strategy

Phase	Scope and Objective
Phase 1 — Pilot	Single product family, single document type. Validate architecture, data model, and template engine.
Phase 2 — Expansion	Extend to 3–5 product families. Establish template governance and data quality remediation.
Phase 3 — Scale	Full portfolio onboarding. Activate localization, multi-market regulatory variants, multi-system sourcing.
Phase 4 — Intelligence	Trigger-driven automated generation, analytics on document usage, and service-stage personalization.

Change Management

Technical writers and documentation specialists must be repositioned from document authors to template designers and data quality stewards — roles offering greater strategic value. Resistance to this repositioning is the most consistent cause of adoption failure in DDP programs.

15. Conclusion

Digital Document Publishing represents a fundamental rearchitecting of how product knowledge is materialized, managed, and distributed across the industrial enterprise. It connects the Digital Thread to the human and regulatory audiences who need its content, in the format they require, at the moment it is needed.

The HVAC case study documented here —months to hours, zero IT involvement for template management, multi-format output, governed approval workflow, PDF/A archival — is a replicable outcome. The platform architecture is domain-agnostic and document-type-agnostic. Its implementation pattern scales from a single product family in a single plant to a global portfolio across multiple regulated industries.

With proven DDP foundation, the adoption of autonomous AI and agentic DDP marks the point where documents cease to be static computed artifacts and instead become intelligent, living projections of the product knowledge base with latest information—constructed with contextual reasoning, proactively validated, and continuously evolved by AI agents that learn from every interaction.

Because the architecture is PLM-agnostic, this capability is available today across Teamcenter, Windchill, 3DEXPERIENCE, Aras, or any platform—without migration, without lock-in, and without displacing existing investments.

STRATEGIC RECOMMENDATION	Treat your document library as a real-time projection of your engineering system of record. Every document manually authored is a document that will drift from the truth. Every manual approval email is an audit gap. Every locale variant maintained separately is a multiplication of risk. Digital Document Publishing ends all three — architecturally, not procedurally.

15. References and Standards

International Standards

- ISO 19005-1/2/3 — Electronic document file format for long-term preservation (PDF/A-1, -2, -3)
- ISO 9001:2015 — Quality Management Systems | ISO 13485:2016 — Medical Devices QMS
- IEC 61508:2010 — Functional Safety | IEC 62304:2015 — Medical Device Software

Industry Standards

- AS9100D — Aviation, Space & Defence QMS | IATF 16949:2016 — Automotive QMS
- FDA 21 CFR Part 820 / Part 11 — Medical Device Quality & Electronic Records
- EU MDR 2017/745 | ATEX Directive 2014/34/EU | EU REACH Regulation (EC) No 1907/2006

Technology References

- Aras PLM R22/R37 — REST API Developer Documentation, Aras Corporation
- Aspose.Words for .NET — Document Generation and Rendering Library, Aspose Pty Ltd
- veraPDF 1.x — Open Source PDF/A Validation Library (ISO-endorsed), The veraPDF Consortium
- Infosys Engineering Services — Technical Documentation Automation Framework, Internal Reference Architecture

About the Author

Shirish Topiwala
Technology Architect, GERDLY, Infosys

Sibaji Pattanaik
Principal Consultant, IOT, Infosys

Kapil Manikrao Jadhav
Industry Principal, IOT, Infosys

For more information, contact askus@infosys.com

Infosys®
Navigate your next