How SAP Joule redefines the SAP experience

WHITEPAPER

ust.com

UST

Table of contents

1. Executive summary 03
2. Introduction: defining the conversational enterprise 04
3. Joule's operational framework 05
4. Transformational impact across the SAP suite 06
5. Technical deep dive: architecture, grounding, and security 07
6. Core business benefits and ROI 12
7. Implementation, governance, and best practices 13
8. Conclusion 14
9. Next steps 15

1. Executive summary

SAP Joule marks a meaningful shift beyond traditional robotic process automation and basic chatbot solutions. It operates as a unified generative AI copilot embedded across the SAP cloud portfolio, helping organizations access and act on complex enterprise data through natural language interactions. By running on the secure and scalable SAP Business Technology Platform, Joule goes beyond task assistance. It functions as a context-aware agent that can carry out multi-step business processes, surface real-time insights from analytics, and support developers in working more efficiently.

This paper takes a closer look at the technical underpinnings of SAP Joule, explores how it is reshaping core business workflows, and outlines the strategic value organizations can expect to gain. It also examines potential return on investment and shares practical guidance for organizations planning to implement Joule successfully.

2. Introduction: Defining the conversational enterprise

Generative AI's ability to interpret user intent and respond in natural, human-like language represents a clear change in how people interact with enterprise software. Before Joule, SAP users worked within highly structured, system-driven interfaces. Completing even straightforward tasks often required navigating multiple applications, learning transaction codes, and following rigid workflows across complex menus. This approach demanded significant time and mental effort, created barriers for many business users, and slowed both adoption and decision-making.

Joule removes much of this friction by shifting enterprise interaction from task execution to outcome-oriented dialogue. Rather than moving through a series of technical steps across Fiori applications and reports, users can state a business goal in plain language, such as "Improve my cash position." Joule interprets the intent behind the request, identifies the relevant data and processes, and surfaces insights or actions through a unified conversational experience.

By securely bringing together structured SAP transactional data and unstructured enterprise knowledge, Joule lowers cognitive burden, improves accessibility, and helps organizations move from insight to action more quickly. The result is an SAP experience that feels more intuitive, more inclusive, and more focused on enabling productivity across the enterprise.

3. Joule's operational framework

Joule is designed to deliver a consistent and intuitive experience, regardless of which SAP application it engages with.

3.1 The four pillars of interaction

Joule's interactions are organized based on the level of system access and action they require:

Pillar	Action scope	Specific example	System impact
TRANSACTIONAL	Executes system changes (Create, Update, Delete).	Create a sales order for Acme Corp for 50 units of Product X, standard pricing.	Directly updates SAP S/4HANA tables.
ANALYTICAL	Extracts, summarizes, and visualizes complex data.	Compare margin trends for the last two quarters in the APAC region.	Generates visual dashboards and reports on demand.
INFORMATIONAL	Retrieves answers from documents and knowledge graphs.	What are the steps for submitting a cross-border travel request according to the updated HR policy?	References unstructured policy data.
NAVIGATIONAL	Directs the user to deep-link functionality.	Open the 'Maintain Business Partner' app and pre-fill data for vendor #5000.	Accelerates access to specific Fiori apps or screens.

3.2 Mastering the prompt: user interaction

Joule's effectiveness depends on the user's ability to clearly articulate intent and provide sufficient context. To support this, Joule applies two foundational concepts. The first is **Contextual Awareness**, which incorporates factors such as user identity, location, and prior conversation history. The second is **Chain-of-Thought**, which enables Joule to break complex requests into a series of logical, executable steps.

4. Transformational impact across the SAP suite

Value is delivered by reducing manual effort and extending expertise across critical business functions.

- **ERP (S/4HANA)**: Automates intercompany account reconciliation, supports natural language queries for predictive analysis such as supply chain disruption scenarios, and identifies potential financial anomalies.
- **HXM (SuccessFactors)**: Produces candidate shortlists, streamlines employee self-service for complex HR events such as relocation, and explains outcomes from automated talent acquisition processes.
- **Development (SAP Build and SAP BTP)**: Speeds low-code development by translating natural language requests into deployed CAP service logic and offering context-aware support for debugging and issue resolution.

5. Technical deep dive: Architecture, grounding, and security

Joule is built as a cloud-native, enterprise-grade generative AI copilot on the SAP Business Technology Platform. It is designed to deliver deterministic, auditable, and context-aware capabilities across mission-critical SAP applications.
The underlying architecture is organized into three planes: control, data, and execution. Together, these layers support security, governance, and operational reliability at enterprise scale.

1. Control plane: Governance, model orchestration, and policy enforcement

Within the control plane, Joule relies on the SAP BTP AI Foundation as the central layer for governance and orchestration. This plane is responsible for managing how AI capabilities are selected, secured, and monitored across the platform.

- **Model abstraction and orchestration**: The Generative AI Hub routes user requests to SAP-approved large language models, including OpenAI GPT-4, Anthropic Claude, and Gemini. Model selection is governed by policy and informed by factors such as contextual relevance, latency requirements, token usage, and enterprise compliance constraints.
- **RBAC and identity propagation**: Joule integrates with SAP Cloud Identity Services to enforce tenant isolation, role-based access controls, and single sign-on across all supported endpoints.
- **Prompt governance**: The platform supports versioned prompt templates, controlled context injection, token limits, and content filtering. These safeguards help ensure predictable behavior and alignment with enterprise governance requirements.
- **Audit and compliance logging**: Every interaction is recorded, including the context retrieved, the model version used, and any actions triggered. This provides full traceability and supports compliance with regulatory frameworks such as GDPR, CCPA, and SOX.
- **Skill lifecycle management**: Through Joule Studio, developers can design, test, deploy, and monitor custom Joule skills. The environment supports both low-code and pro-code development, enabling extensibility while maintaining governance standards.

2. Data plane: Secure context retrieval and vectorized knowledge access

Joule uses a data-grounded architecture to ensure that every AI response is both traceable and grounded in the enterprise context. This approach is designed to maintain accuracy, transparency, and trust when working with sensitive business data.

- **Ephemeral, scoped access**: Enterprise data is accessed on a just-in-time basis using temporary credentials. Data does not persist within the model layer, reducing exposure and limiting risk.
- **Structured and unstructured data integration**: Joule retrieves information from SAP HANA tables and CDS views, as well as from unstructured sources such as documents, PDFs, and internal knowledge repositories.
- **Vectorization and semantic search**: Relevant data is converted into dense vector embeddings and indexed in SAP HANA or dedicated vector stores. This enables efficient similarity search during retrieval-augmented generation workflows.
- **Semantic layer and knowledge graph**: Technical SAP artifacts, such as MARA-MATNR and BKPF-BELNR, are mapped to business-level concepts. Knowledge graphs capture entity relationships, hierarchies, and business rules, supporting multi-step reasoning and reliable query translation.
- **Data redaction and masking**: Sensitive fields are masked or tokenized in accordance with SAP authorization policies, helping ensure regulatory compliance and protection of sensitive information.

3. Execution plane: Retrieval-Augmented Generation (RAG) and action framework

Joule translates AI-generated insights into reliable business outcomes through retrieval-augmented reasoning and governed execution. This execution layer ensures that responses are not only informative but also actionable within enterprise systems.

- **Semantic parsing and intent resolution**: User requests are analysed to identify intents, entities, and relevant contextual parameters.
- **Contextual retrieval**: Pertinent structured and unstructured data is retrieved, ranked, and normalized to ensure high-quality context is provided for prompt construction.
- **LLM inference**: An SAP-approved large language model generates responses grounded in enterprise data. This approach supports accuracy, traceability, and explainability.
- **Action framework and transaction safety**: When appropriate, responses can initiate API or OData operations. These actions are subject to pre-execution validation, simulation, and optional human review, helping ensure safe execution in sensitive domains such as finance, human resources, and supply chain operations.
- **Observability and telemetry**: Key metrics, including latency, token consumption, retrieval relevance, and execution outcomes, are continuously monitored to support optimization efforts and adherence to service-level expectations.

4. Developer extensibility and skill deployment

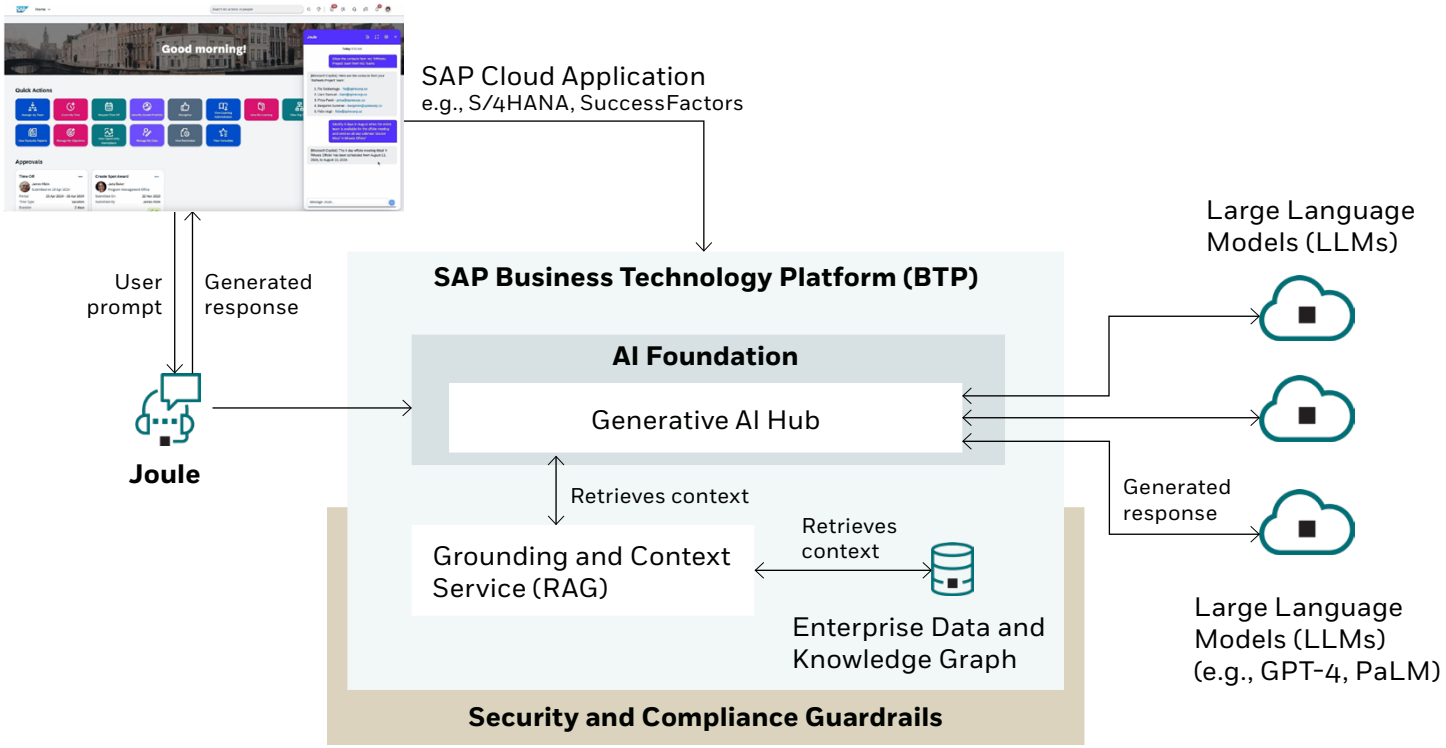
- **Joule Studio**: Provides a governed environment for creating and deploying custom AI skills that integrate directly with SAP workflows, including ERP, human resources, and analytics modules.
- **Reusability and composability**: Skills can be designed to coordinate multiple large language model interactions, invoke APIs, and apply deterministic business rules, enabling modular and repeatable automation patterns.
- **Performance metrics**: In observed deployments, organizations have reported approximately a 25 percent reduction in manual task effort and up to a 10 percent improvement in skill deployment speed when compared with earlier generations of AI-driven automation tools.

5. Enterprise scale and governance

Joule is designed to support multi-tenant, large-scale deployments and can serve tens of thousands of enterprise users within a single environment. Looking ahead, SAP has outlined plans to expand Joule to support more than 400 AI use cases and thousands of skills by 2025. These capabilities are intended to deliver auditable, context-aware AI that is embedded directly into both operational and strategic business workflows.

6. Technical flow overview

1. **User query**: User submits a request through Fiori, Work Zone, SAP GUI, or an API.
2. **Control-plane authentication**: The system enforces role-based access controls and propagates user identities.
3. **Semantic parsing**: Joule resolves business intent into entities, parameters, and actionable steps.
4. **RAG retrieval**: The platform retrieves and ranks relevant structured and unstructured context.
5. **Prompt construction**: The system injects validated context into model-ready LLM prompts.
6. **LLM inference**: An SAP-approved model generates deterministic, data-grounded responses.
7. **Optional action execution**: The platform executes API or OData operations with validation and optional human approval.
8. **Observability and audit**: The system captures metrics and logs to support compliance, performance monitoring, and continuous optimization.



6. Core business benefits and ROI

Adopting Joule delivers measurable strategic and operational value, improving return on investment across the enterprise.

Benefit category	Description	Impact
Productivity and efficiency	Automates repetitive transactional and informational activities, reducing time spent navigating systems and retrieving data.	**Significant reduction** in time spent on common tasks such as approvals and data lookups.
Decision quality	Delivers immediate, context-aware analytical insights by synthesizing information across multiple systems and data sources.	**Improved forecast** accuracy supported by faster anomaly detection.
User experience (UX)	Reduces the learning burden associated with complex Fiori applications and layered menu structures, enabling faster user adoption.	**Increase in system adoption rates** paired with a reduction in user training effort and costs.
Data governance and compliance	Ensures that all AI-initiated transactions comply with predefined security authorizations and remain fully auditable within core systems.	**Higher compliance adherence** and a corresponding reduction in the risk of manual security errors.
Developer velocity	Accelerates the development of new applications and extensions through text-to-code generation, improving developer productivity and time to value.	**Faster prototyping and extension of SAP applications,** accelerating development timelines.

7. Implementation, governance, and best practices

Successful adoption of Joule depends on a proactive approach that addresses security, ethical considerations, and organizational change management.

7.1 Responsible AI and ethical guardrails

SAP enforces its commitment to Responsible AI through the SAP Business Technology Platform.

- **Data isolation**: Customer data remains strictly segregated and is not used to train third-party foundational large language models.
- **Auditability**: Every action executed by Joule is logged and fully traceable, and is subject to the same compliance controls and audit requirements as actions performed manually by users.

7.2 Change management and adoption

The focus shifts from training users to navigate menus to training them in prompt engineering, helping them learn how to ask effective questions and fully leverage Joule's contextual understanding.

7.3 Joule adoption: do's and don'ts

DO'S

Prioritize clean data

Ensure that core master data, including customers, products, and vendors, is accurate and consistent. Joule depends on this foundation to deliver reliable, data-grounded responses.

Define the semantic layer

Actively refine the mapping between technical SAP data structures and the business terminology used across your organization. Clear semantic alignment improves intent resolution and response accuracy.

Focus on workflow augmentation

Apply Joule to complex, multi-step workflows that currently require users to navigate across multiple Fiori applications. These scenarios often deliver the highest productivity gains.

Start with high-volume informational queries

Begin with use cases such as HR policy lookups or documentation retrieval, where efficiency improvements are easy to measure and value can be demonstrated quickly.

DON'TS

Do not expect perfect answers on day one

Joule improves as it is tuned to your organization's terminology, workflows, and usage patterns. Plan for iterative refinement, structured feedback, and ongoing optimization.

Do not deploy in isolation

Joule delivers the most value when it is integrated across functions and data domains. Avoid limiting early deployments to a single area, such as finance, if cross-functional context is required for meaningful outcomes.

Do not skip governance

Establish clear security roles, audit requirements, and approval workflows upfront. All Joule-driven actions must operate within existing user authorizations and compliance controls.

Do not underestimate change management

Adoption is not automatic. Users need enablement and reinforcement, including practical training on prompt crafting, to achieve sustained usage and measurable return on investment.

8. Conclusion

Joule brings together deep business process expertise, rich enterprise data, and advances in generative AI into a single, integrated capability. It reflects a shift in how users interact with enterprise systems toward more natural, intent-driven engagement. This change supports greater efficiency, faster and more informed decision-making, and a more intuitive experience across the SAP landscape. Investing in Joule adoption is an investment in the long-term competitiveness and resilience of the Intelligent Enterprise.

9. Next steps

Start with an AI readiness assessment

Evaluate data quality, process maturity, and organizational readiness to identify high-impact opportunities for Joule adoption.

Run a focused Joule pilot

Begin with one or two business-critical processes, such as finance, supply chain, or human resources, to validate value and inform scaling decisions.

Align with SAP Activate and AI enhancements

Use SAP Activate, including its AI-related extensions, to support a structured and repeatable rollout across the enterprise.

Establish an AI governance framework

Define clear policies for security, compliance, adoption, and risk management to support sustainable, enterprise-wide use.

Position for long-term advantage

Organizations that adopt Joule with discipline and intent are better positioned to advance SAP's vision of the Intelligent Enterprise.

Redefine the SAP experience with Joule.

Discover the architecture, benefits, and best practices shaping the conversational enterprise.

Author **Guruprasanna T K**
Architect I - Enterprise Solutions
sap@ust.com →

Talk to an SAP Joule expert →

Since 1999, UST has worked side by side with the world's best companies to make a powerful impact through transformation. Powered by technology, inspired by people, and led by our purpose, we partner with our clients from design to operation. Our digital solutions, proprietary platforms, engineering, R&D, products, and innovation ecosystem turn core challenges into impactful, disruptive solutions. With deep industry knowledge and a future-ready mindset, we infuse expertise, innovation, and agility into our clients' organizations—delivering measurable value and positive lasting change for them, their customers, and communities around the world. Together, with 30,000+ employees in 30+ countries, we build for boundless impact—touching billions of lives in the process.

ust.com

Version 0103-20260220