heuristiq AI Studio × CTG Consulting

AI ADOPTION IN MID-SIZED COMPANIES · ONLINE DIALOGUE

Sovereign AI

Model Choice and *Cost Control*

Why a single model is not enough, and how three levers secure data sovereignty, flexibility to switch and predictable costs: choose, route, control. Your key takeaways.

Issue 3

5 November 2026 | Peter Kramer and Simon Kozlik

AI ADOPTION IN MID-SIZED COMPANIES · ONLINE DIALOGUE | KEY TAKEAWAYS

Sovereign AI: choose, route, *control*

1. **Choose: classification before performance.**
   First clarify what a model is allowed to see, then pick the smallest model class per task that is reliably sufficient. In the example, five of six steps run in Germany or the EU.

2. **Route: distribute requests instead of using the strongest model for all.**
   Each request is assigned by data class, task and confidence. In the example, only around 5 percent need the strongest model, and model costs fall by around 90 percent.

3. **Control: an abstraction layer turns switching models into configuration.**
   Applications never talk to a model directly. Routing, budget, logging and approval are managed centrally. Switching models then takes one to two days instead of three to six months.

4. **Budget per value stream, cost per case, limit per agent.**
   Three KPIs make AI costs manageable, tracked in the value stream rather than in the IT cost center. A hard cost limit per agent prevents loops from silently burning money.

5. **IT runs the abstraction layer, the value stream owner sets the rules.**
   Sovereignty is a leadership decision, not an IT project. Who decides on models, budgets and data must be settled before scaling up.

6. **Cheaper per token, more expensive overall: agents drive consumption.**
   Costs per token are falling, yet spending on generative AI more than tripled in 2025. With agents, costs are shifting from licenses to usage.

The best model is the one you can switch at any time.

Which model will you switch first?

We would be happy to discuss your individual questions in a personal conversation.

Peter Kramer
CTG CONSULTING
Business Transformation & AI Scaling
peter.kramer@ctg-consulting.com
LinkedIn linkedin.com/in/peter-kramer-029728191

Simon Kozlik
HEURISTIQ AI STUDIO
Data, AI & Technical Implementation
simon.kozlik@heuristiq.co
LinkedIn linkedin.com/in/simon-kozlik-0627082a