

Automating yourself out of the job?

Miroslav Jakúbek, Head of Data Protection & DPO, Solaris SE
Paulina Kostrzewska, Head of Privacy Operations & AI Governance, Omnissa

Disclaimer

The views expressed are our own, shaped by our individual perspective on the current industry positions and developments. They do not reflect the views of our current or past employers or clients.

This presentation is for information only and does not constitute a legal advice.

What this talk is about

In scope

AI agents running privacy operations

- DPIAs and privacy reviews
- Intake and drafting
- Review and sign-off controls
- Documentation and tough conversations to be had

Out of scope

Privacy of AI (Agents) processing itself

- Training data and lawful basis
- Security/Lawfulness of AI Agents
- Duties of AI providers
- General (DP) Compliance of the models / agents

The question for today

What must be true before a privacy team lets an agent do part of its own job?

Views my own. Not legal advice.

Why we built agents

Bank: DPIAs drafted in Gemini

Colleagues started to use Gemini to answer DPIAs + clear assignment from C-Level to automate

Answers came out long and wordy, which made DPIAs harder to read and review (**adding workload for DP instead of reducing**)

Doing nothing was not neutral: AI was already shaping DPIA content without controls

Omnissa: slow, heavy reviews

Slow reviews; requesters struggled to provide quality information

Tailored, risk-based reviews still failed: low privacy acumen in the business

Colleagues used AI without grasping nuances; tools gave inconsistent risk advice, losing the company context

Same conclusion

A governed agent beats unguided AI use

Four levels of considerations

Operations

- How does this fit into my general Data Protection Management System?
- How do I deal with the established DPIA/DSR process and a software vendor not supporting (my) AI agents?
- Do I actually i**ncrease** workload by using one?

Policy

- How do I address the use of an AI agent in my policies and procedures?
- (As a bank), I need controls: do I manually review everything to be able to sign off?
- What is the point of an AI agent then?
- What about token usage ? (AI Policy + costs!)

Conflict of interest

- If I co-design an AI agent or a prompt used by departments, am I reviewing my own work? Am I conflicted then?
- Do I need an independent reviewer for anything I helped design?
- How does it differ from designing a PIA/DPIA template?

Technical

- Where do I develop it? (Python script problem) device.
- How do I make sure the agent will not be changed by people using it?
- How do I address people using various AI tools in their daily business?

How we approached them

Operations

- Agent is just another "team member" with limited trust and privileges.
- Agent researches, validates answers & drafts, hands off to human in the loop for confirmation & approval.
- Fix the process first, don't let the agent create its own process;
- Set the guardrails to prevent AI from straying away from the why & the how
- Future proof your agent (time tracking/token cost!)

Policy

- Policy is the what and why; the "how" lives in a process documentation. (or procedure if you are a regulated industry)
- Outputs are labelled drafts, validated by a human: 80% confidence for 20% of the effort.
- Gating, risk-based approval, and layered/tiered approach are your weapons to cut the low risk, low impact reviews, with agent's approval & PIA record

Conflict of interest

- Document everything properly, assign responsibilities (Art 38/39 GDPR!!) and keep the agent scope well-defined, and transparent
- Version the skill, log changes to the templates, design AI Agent with the target audience's input.
- For additional confidence, keep builder, requester and approver roles apart.

Technical

- The tool follows what is permitted & available.
- No one technology that rules them all; stay nimble and adapt. Build a universal foundational agent blueprint easily adaptable to any technology.
- Meet the business requestors where they are at. You won't win this battle, and even more so the war. Soon, every tool will have AI, and companies will use multiple GenAI technologies.

From questions to practice

1. **Operations**
   See: Review process with the AI assistants

2. **Policy**
   See: Governance and accountability, appendix

3. **Conflict of interest**
   See: Human in the loop, I validate

4. **Technical**
   See: My 2026 journey, Copilot to Claude to Rovo

From the legal view to the practical view

How the agents were built, tuned and rolled out, and what changed along the way

My 2026 journey

Microsoft Copilot

Feb – Apr

Tried an agent. My learning: I lacked the expertise, tooling had serious limitations, and agent's performance was low

Claude

Jun – ongoing

Learned, then built 6 skills, continually improved, promoted adoption across teams and stakeholders. Now, about 23% reviews are done with AI or by AI at a pre-screening stage.

Atlassian Rovo

Aug – ongoing

First agent built and tested; now in beta with a small user group.

Views my own. Not legal advice.

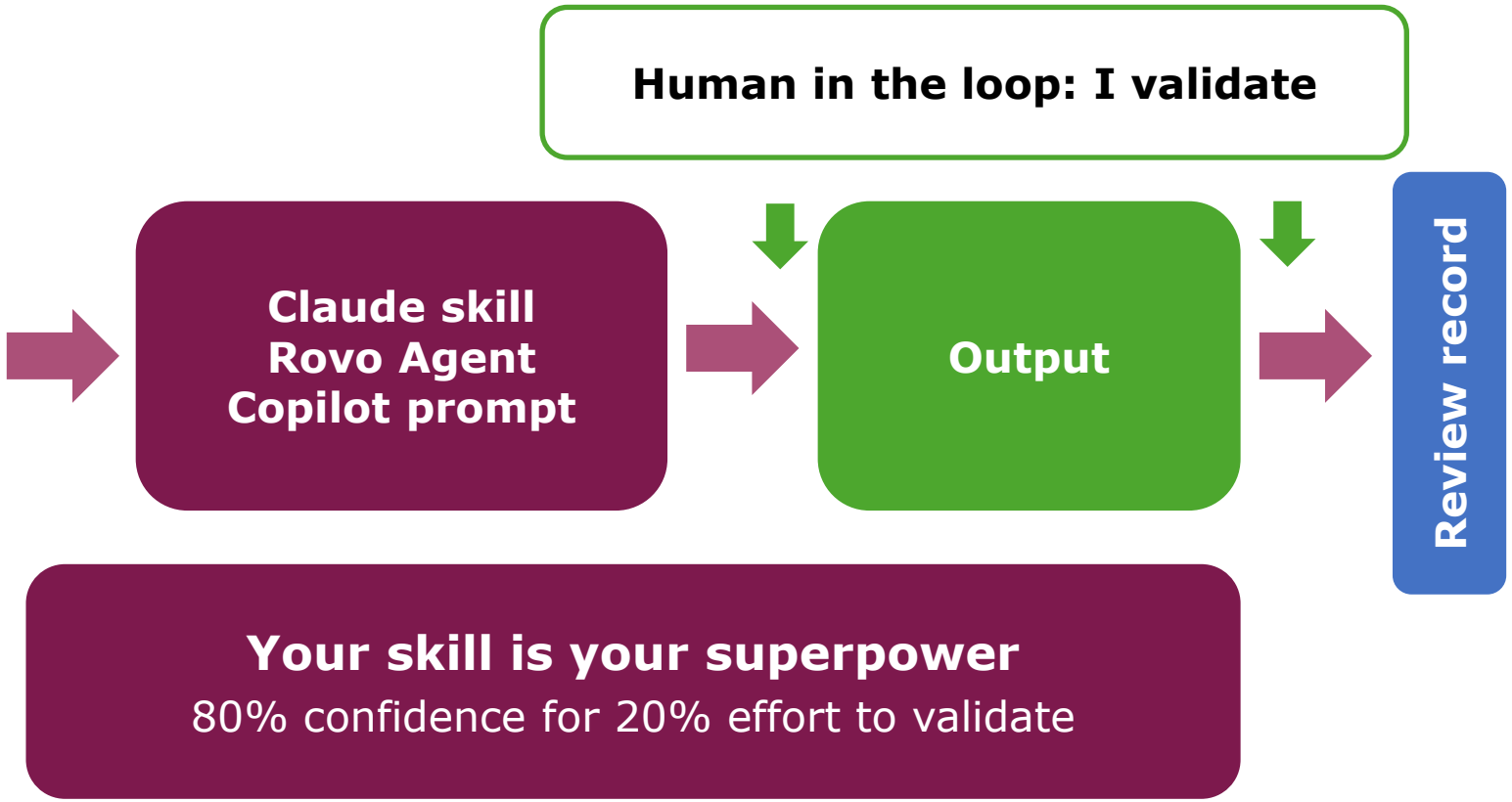
Learning Claude, month by month

Month 1	Month 2	Month 3+
June Simple tasks, training, Claude 101, trial and error. Lots of tokens spent	**July** First skill: I described the review ecosystem, goal, laws, policies and desired output. I treated Claude as a new team member. All conversational, knowledge transfer approach to build a skill.	**August & ongoing** 6 skills, continually improved, with growing adoption.

How I approached it

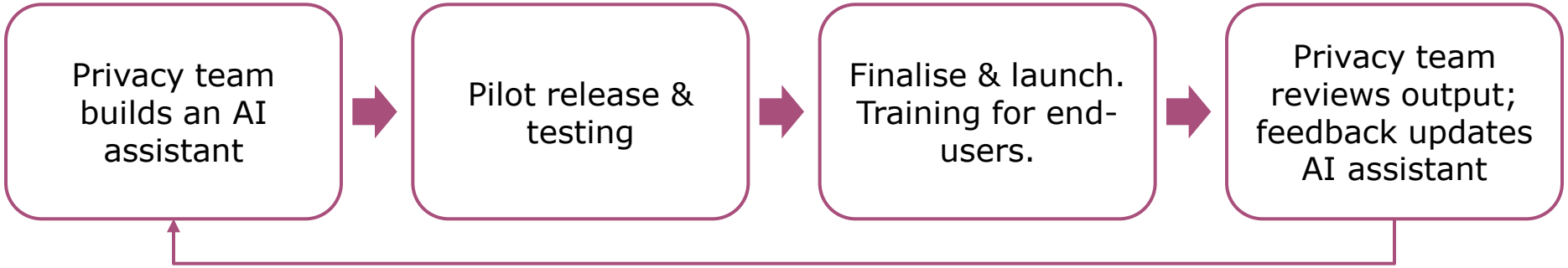
Programmed on

- Company policy & contract positions
- Privacy programme baseline
- Laws & frameworks
- Controller vs processor role
- Desired output structure and contents
- Human-in-the-loop steps
- Gates, risk, auto-approval tiers
- Other….

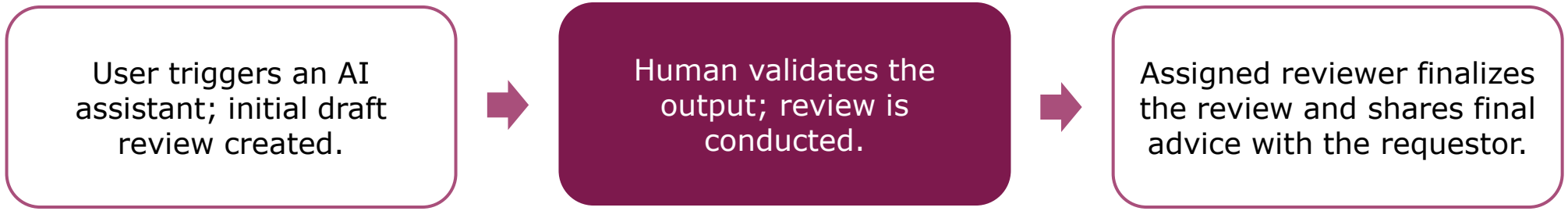


Review process with the AI assistants

Build and tune prompts / agents / skills



Review process



Lessons from the DPIA agent at the bank

"I never once failed at making a light bulb. I just found out 99 ways not to make one." Thomas A. Edison.

1. **Solve "boring" questions first**
   Amendments to policies and produceres, compliance with AI governance policy (token usage!!), target audience (MAC vs Windows vs Linux)

2. **KYS (Know your stakeholders)**
   Developers are unlikely to understand the legal framework that the AI Agent should consider (catch it all approach fails hard)

3. **What exactly are you automating? Be precise!**
   General DPIA AI Agent is going to fail. Finding a precise part of DPIA (i.e. creating data flow from existing resources) is much more likely to succeed

4. **Set realistic targets**
   Be aware of your capacities, team capacities and developer capacities. Currently 5th version of DPIA AI agent hat to be discarded due to various issues (**short deadline = long days and short nights**!)

Tough conversations checklist

- ❑ You have AI Agents, can we cut headcount? (**C-Level**)
- ❑ How come there is a backlog if you have AI Agents? (**C-Level/ Supervisory board**)
- ❑ Am I training my own replacement ? (**direct report**)
- ❑ I copy pasted what the AI (agent) said, you are checking it anyway ? (**business stakeholder**)
- ❑ Who is solving all the “Adjacent boring problems”?

Possible solutions:

- ✓ Have numbers ready – real time savings vs imaginatory AI Agent utopia
- ✓ Reference to policy (that was properly amended!) + numbers
- ✓ Have honest conversation and focus on short term gains and upskilling
- ✓ Make sure responsibilities are properly assigned, understood and documented in policies/procedures
- ✓ Make sure you do not rely on other people to do the work nobody wants to do

Key takeaways

1. **Change management is key**
   Keep iterating to drive the best outcomes
2. **Knowledge into steps**
   Turning legal knowledge into actionable steps takes time
3. **Process readiness**
   AI can improve a bad process, but won't build a good one from scratch
4. **New talent needs**
   Close knowledge gaps to improve adoption
5. **Forest and trees**
   Keep the review goal in view while fixing the details
6. **Prioritize relentlessly**
   Many specialized agents beat one big, complex agent
7. **Measure impact, monitor for drift**
   Keep tracking results so accuracy does not slip over time

0. Technology prerequisites

Best practices for a privacy-review skill

- **Understand the technology you want to use:** how it works, what it can do, what is possible within your organization (e.g., system integrations)
- **Let the technology teach you first:** e.g., use /skill-builder on Claude, Agent builder on Rovo and similar to build your first Ai assistant. Ask Claude to be your guide on building the skill; reuse your AI knowledge across the systems.
- **Don’t let the technology limit your thinking**: rather than assuming something isn’t possible, ask the technology how to execute it
- **Build one blueprint, adapt to technology as needed:** don’t build from scratch for any new AI technology. Your process is your main driver. Clearly defined gates, AI-to-human handover or gating, repeatable outputs are within your control and scale better, and faster. Well designed blueprint is the key to quick agent scale up on the next big and trendy technology that comes.

Appendix 1 – Practical observations

1. Scope and routing

Best practices for a privacy-review skill

- **Be crystal clear on your scope:** the easiest way to fail is to poorly define what the AI assistant is there to do and risk introducing scope creep
- **Define one job per skill:** screening, full DPIA, vendor assessment, gap analysis and service-level review are separate skills. Bite-size knowledge build than you can later bundle together, more flexibly. Easier maintenance in future
- **State the boundary in the description:** what it does, what triggers it, which skill to use instead
- **Write for a named audience:** engineers, business users, privacy specialists and auditors need different language and depth
- **Use tiers:** a light screening that escalates to a full review is easier to adopt

2. Intake and autonomy

Best practices for a privacy-review skill

- **Collect input once, up front:** run autonomously, then do a single final review
- **Classify every field:** inferable (default it, tag as assumption), researchable (cite a confirmed source), human-only (ask only for these)
- **Prefer structured choices** over open questions, and always leave an "Other" option
- **Confirm the source before researching:** the user confirms the official URL or document first
- ***Check for prior work:** update earlier reviews of the same subject instead of duplicating them (*depending on skill configuration, or adoption)

3. Accuracy and anti-fabrication

Best practices for a privacy-review skill

- **Never invent facts:** use visible markers: Data unavailable, [ASSUMPTION: …], [TO CONFIRM: …]
- **Tag citations by reliability:** pinpoint articles, effective dates and guidance are where fabrication is most likely
- **No silent supplement:** if research is thin, report it instead of filling the gap from model memory
- **Don't hard-code contact details** you can't verify
- **When unsure, score at the higher risk tier** and record why

4. Review methodology

Best practices for a privacy-review skill

- **Anchor to a recognized framework** and reuse the same scoring scale across all skills
- **Build an internal policy screening step**: check the reviewed activity for compliance with internal policies. Map the gaps, add it to final risk report
- **Make rules explicit:** escalation thresholds, mandatory-review indicators, prohibited-practice screens, risk-to-action mappings
- **Apply the strictest requirement** across jurisdictions and flag where laws diverge
- **Model roles correctly:** controller vs processor under each applicable law
- **Review contracts as well as law:** DPA terms, sub-processor commitments, notification duties
- **Keep a data-flow diagram as an image;** give every gap a mitigation, an owner and a date

5. Governance and accountability

Best practices for a privacy-review skill

- **Label outputs as drafts or provisional:** not a legal opinion or an approval, unless specifically designed for the skill (e.g., pre-screening vs full DPIA)
- **Add a notice to every output:** "Draft only. Review required by Legal, Compliance, or the relevant control owner."
- **Humans own declarations and sign-off:** the skill never signs itself off
- **Match desired actions to risk levels:** low risk gets a record only, higher risk gets escalation; define auto-approval narrowly
- **Redact personal data and secrets,** and tell users to rotate any secret that was shared

6. Outputs and integrations

Best practices for a privacy-review skill

- **Treat the output as a system of record:** stable structure, consistent field names, version history, finding IDs that are never reused
- **Update in place for re-reviews** and keep an audit trail
- **Treat field names as a contract** if dashboards or registers read them
- **Link to authoritative sources** instead of copying them, so the review doesn't go stale
- **Document connector limits and fallbacks,** such as no file attachments or email, work with your AI technology to help you figure it out
- **Verify before claiming success:** never say something was created, sent or applied unless checked

7. Maintenance

Best practices for a privacy-review skill

- **Version the skill** with semantic versioning and a changelog, kept in one authoritative place
- **Keep shared logic consistent** across skills: thresholds, scoring and policy baselines
- **Check for contradictions** in instructions on sending, approving or ordering steps, and that every referenced file resolves
- **Avoid hard-coded names, IDs and page numbers;** they break silently when people and structures change
- **Schedule a policy and law refresh** so triggers and baselines stay current
- **Test with real scenarios,** including missing information, conflicting sources and high-risk inputs
- **Invest time in reviewing user feedback**, your skill is worth nothing without its users
- **Communicate back what you are fixing and why**, transparency with users builds trust and repeatable process strengthens adoption