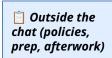


Responsible Al Roadmap

Risk & Quick Mitigations (Your 80/20 Checklist)
Navigate Al. Advance what matters.



v2025.09.04 latest version at aligniq.ai/roadmap



Inside the chat (prompting practices)

Step 1 Sandbox → Basics to experiment safely

(Tildentify Safe vs. Unsafe Data

- ☐ Use company- approved tools
- ☐ Turn off training in public chats
- ☐ Classify Red / Amber / Green

Pro tip: Mask names/IDs on Amber data

Data traffic light defaults (org policy overrides)

Red = secrets/regulated → never paste.

Amber = internal/drafts → mask & minimize.

Green = public→ safe. When unsure, treat as Amber

☐ Never paste Red

Expect & Detect AI Hallucinations

- ☐ Follow the 3-source rule ☐ verify numbers & dates Pro tip: Run the "could I defend
- **Pro tip:** Run the "could I defend this?" test: rewrite it, flag what you can't defend, and verify those parts
- Ask for citations
- Request a brief counter-argument

SMaintain Stakeholder Trust

- Disclose Al editing, summarizing, and drafting help
- ☐ Don't forward raw Al output
- ☐ Keep **human sign-off** and mention it
- ☐ Keep an experiment log (prompt + date + purpose) for work that impacts stakeholders.

Pro tip: Keep a disclosure template bank, and match details to stake and sensitivity.

Step 2 Workflow → Daily responsible use

Protect Private Data at Work

- Use your work account
- ☐ Ensure retention matches IT policy
- ☐ Work only in company-account Al tools & connectors.

Pro tip: *Test new prompts/tools with de-identified snippets first.*

- Prioritize & Mitigate Bias
- ☐ **Triage** risks by likelihood and severity
- Review outputs across multiple personas
- ☐ Run persona-swap tests
- □ Add inclusive constraints

Pro tip: Ask for plain language, alt-text/captions, non-color cues before finalizing.

- **Keep Healthy AI Boundaries**
- ☐ Keep a never-delegate list;☐ draft your POV first.
- Follow the 3-source rule
- verify numbers & dates

Pro tip: Schedule weekly no-Al reps / clarity breaks.

- ☐ Have Al critique vour draft
- Have Al ask you clarifying questions.

∰ Step 3 Strategy → Advanced/long-term alignment

Navigate Access & Compliance

- Confirm lawful basis & data residency
- ☐ Apply least privilege access

 Pro tip: Regularly review Al use
 and retire what doesn't work.
- ☐ If unsure, treat it as **RED** and stop
- □ Report issues (bias, leak) to Al lead/IT

- SAlign AI with Mission & Values
- ☐ Keep do/don't-automate list.
- ☐ Start projects with a **values preamble** + audience in mind.

Pro tip: Log when you choose not to delegate to AI and why.

□ Inject values into prompts

- ₩Use AI Sustainably
- Choose right-sized models for the task
- Ask vendors for energy, water, and carbon disclosures

Pro tip: Apply at scale for org-level usage not individual prompts

- Prefer textover media
- Batch prompts