Responsible AI at Work

AI tools can draft emails, summarise reports and analyse data in seconds, and many of us now use them every day. But speed is not the same as responsibility. The OECD AI Principles, agreed by many governments, and the AI Risk Management Framework of NIST, the US standards institute, describe what trustworthy AI looks like: it respects privacy, it is accurate and fair, it is transparent, and people stay in control. In this title, we turn these ideas into five practical habits for everyday work. Each habit comes with simple do's and don'ts.

1. Protect Privacy

The first habit is to protect personal and confidential information. Anything we type into an AI tool may be stored or used outside our control, depending on the tool and its settings. So we use only the tools that our company has approved, and we remove names, numbers and client details before we paste any text. We never enter passwords, health data, salaries or secret plans into public tools.

Example: A parent wants help writing a letter to a school about a child's illness. To protect privacy, the parent leaves out the child's name and medical details, and the letter is still clear.

At work: Rowan, an HR specialist, needs a summary of a difficult employee complaint. She protects privacy by replacing names with job roles and by using the company's approved AI tool, so no personal data leaves the company.

2. Check Accuracy

The second habit is to check what the AI produces. AI tools can sound very sure of themselves and still be wrong; they may even invent facts, figures or sources. So we check important facts against reliable sources, and we test numbers and calculations. We never send AI output to a client or a manager before reading it carefully ourselves.

Example: A student asks an AI tool for famous sayings to use in an essay. Checking accuracy, the student searches for each saying and finds that one of them was never said by that person.

At work: Rowan asks the AI tool to summarise a new labour law for the managers. She checks accuracy against the official text, finds a wrong notice period, and corrects it before the summary goes out.

3. Watch for Bias

The third habit is to watch for unfair bias. AI learns from data made by people, so it can repeat old patterns, for example about gender, age or nationality. So we review AI suggestions that affect people, such as hiring, pay or performance, and we ask whether every group is treated fairly. We never let an AI tool rank or reject people on its own.

Example: A small online shop asks an AI tool for pictures of its customers, and every engineer in the pictures is a man. Watching for bias, the owner asks for more varied pictures, and the new page speaks to all customers.

This document is a property of Asaseeyat Academy - © Copy Right Protected.

At work: Rowan uses AI to draft a job advertisement. Watching for bias, she notices words such as young and energetic, which can discourage older applicants, so she replaces them with the skills that the job really needs.

4. Be Transparent

The fourth habit is transparency: people should know when AI has played a real part in the work. Being open builds trust and helps others judge the result correctly. So we tell colleagues and clients when AI has drafted or analysed something important, and we follow company rules and copyright when we use AI content. We never present AI work as our own expertise, and we never hide it when someone asks.

Example: A volunteer uses AI to translate an article for a community newsletter. Being transparent, the volunteer adds a short note, so readers know to report any strange wording.

At work: Rowan sends managers a report on a staff survey that AI helped to analyse. Being transparent, she adds a line on how AI was used and which parts she checked, and the managers trust the figures more.

5. Keep a Human in Charge

The fifth habit is human oversight: a person stays responsible for every decision that matters. AI can suggest, sort and draft, but it cannot be held accountable; people can. So we decide in advance which tasks need a human review, and we keep the final say on decisions about people, money, safety and the law. We never let AI decide alone when the result could seriously affect someone.

Example: A navigation app suggests a road through a flooded area. Keeping a human in charge, the driver looks at the road, ignores the suggestion and takes a safer route.

At work: An AI tool suggests rejecting a candidate because of a gap in the CV. Rowan keeps a human in charge: she reads the application herself, learns that the gap was a year of caring for a parent, and invites the candidate to an interview.

Conclusion

In summary, responsible AI at work rests on five habits: protect privacy, check accuracy, watch for bias, be transparent, and keep a human in charge. These habits put the OECD AI Principles and the NIST framework into practical form. AI is a powerful assistant, but the responsibility for the result always stays with us. Used with care, AI saves time and builds trust instead of putting it at risk.

Sources

- OECD (2019, updated 2024): OECD AI Principles – https://oecd.ai/en/ai-principles
- NIST (2023): Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST AI 100-1 – https://www.nist.gov/itl/ai-risk-management-framework

This document is a property of Asaseeyat Academy - © Copy Right Protected.