



Workshop 1: FORGOOD Updates, Latest AI-Research & Pre-mortem

Summary

FORGOOD Updates

- **FCA and Consumer Duty:** Ran an introductory session on FORGOOD with the FCA. Consumer Duty is their principal lens for AI in financial services, focusing on outcomes rather than regulating AI systems directly. Positive feedback and ongoing dialogue. Goal: FCA familiarity with FORGOOD so members can reference it confidently in regulatory conversations.
- **HMRC and Duty of Candour:** Ran a session with HMRC (UK Revenue & Customs), who are now under requirements to show a duty of candour and explain their thinking and decision-making. They are interested in exploring how FORGOOD can help them be accountable for that. Approaches also received from Defra and other public sector bodies, signalling growing interest across government.
- **Website update:** Members area now structured across workshop recaps, research digests organised by topic (e.g. AI, ethics, organisational bias) with practical practitioner insights and clear FORGOOD action points, and downloadable templates and tools. Digests and resources are the main focus. **Institutional access is included in the membership; please contact us to sign up additional team members.**
- **2026 workshops:** Workshop 2 (11th June): Sludge and administrative burden; Workshop 3 (17th September): Natural Language Processing (AI analysis) of customer complaints in finance. Workshop 4 (3rd December): AI chatbots in consumer and employee contexts. Each session is planned to feature expert speakers.
- **Profile raising:** More visibility planned through LinkedIn, publishing, and conferences.

Latest Research on AI

AI Sycophancy and Recommendations

- AI chatbots are often programmed to validate users rather than challenge them. Even when guardrails are in place (e.g. refusing clearly harmful requests), the way chatbots decline is often so polite and equivocal that users may still interpret it as tacit approval.
- Organisations developing their own customer-facing tools will inherit baseline sycophantic tendencies from underlying models (OpenAI, Gemini, Anthropic etc.), which need to be actively managed.
- LSE study underway examining effects of communication styles in AI financial advisors on accuracy, trust and belief change and what optimal disclaimers from AI look like.
- Literature evolving rapidly across academia, media debate, and regulatory discussion. The FORGOOD team produces ongoing research digests and tools for members.

Cognitive Surrender in AI usage (“Prompt Puppets”), Deskilling and Cognitive Risk

- Key terms discussed: users as “cognitively surrendering prompt puppets” (when users don’t critically question, challenge and engage with AI outputs) and “AI slop” as output (low-quality, uncritical AI-generated output entering workflows).
- Related concepts: “need for cognition” (individual differences in how much people enjoy thinking through problems) and “system three thinking” (the cognitive capacities that direct how you interact with assistive technologies, building on dual processing).
- A particular long-term concern: if AI removes the learning-by-doing that builds expertise, the people who currently define what good looks like may lose those skills over time, leaving organisations unable to improve or govern their AI in future.
- These are not just individual risks but cultural ones: widespread cognitive offloading can erode the quality of collective decision-making.
 - Open question raised: what skills should organisations prioritise in learning agendas, and what competencies and behaviours should they be hiring for given these changes? Communication clarity and technical literacy (i.e. the ability to instruct AI effectively) were identified as key human skills to amplify.
- One member shared that they are framing this internally as “cognitive risk”: the risks to decision quality, tacit knowledge and judgment when AI is integrated into workflows.
 - Challenges raised: who should own such cognitive risk in an organisation?
 - Discussion explored whether it sits under people risk or operational risk, with several members pointing toward joint CHRO/CIO ownership
 - Concerns included untransparent decision chains with different stakeholder involved and negative impacts sometimes only showing up later on, when, erosion of judgment skill in operational teams, employees creating their own AI shortcuts outside standard operating procedures.
- **More FORGOOD tools and research digests coming onto the website about “good versus bad” AI user profiles and cognitive engagement strategies across use cases.**

Agentic AI

- Literature fragmented and immature; lacks practical, actionable guidance on governance. Much of the discourse overpromises on scaling without addressing the factors in between: trust, manipulation risk, privacy, the loss of human interaction.
- The question is not just how to replace a human process with an agentic one, but whether you should. Some organisations have already rolled out and swiftly rolled back AI deployments due to trust and privacy backlash.
- A key insight from financial services: customer barriers are often not informational but emotional (procrastination, anxiety, cognitive difficulty). AI does not solve this.
- As agentic AI develops, employees may increasingly need to relate to AI not just as a tool but as a colleague or even a supervisor, with unintended consequences for working relationships, accountability and organisational culture.

AI Chats as Clean Decision Spaces

- Current AI chat interfaces offer calm, reflective environments for decision-making, which can support more considered choices.
- Risk that these spaces become cluttered with monetisation through advertising, undermining their value as clean decision environments. Particular concern: if users cannot tell whether a recommendation is paid or organic, trust erodes.

FORGOOD Pre-mortem Exercise

The FORGOOD pre-mortem helps surface risks and challenge assumptions before committing to a decision. By imagining a product, service or practice has already failed, participants are freed to identify risks they might otherwise suppress. In the context of the topics discussed, pre-mortems force a pause to actively challenge decisions and risks in these accelerated ai times, imagining failure early to reveal blind spots before they are scaled.

Participants worked in groups through one of two fictional scenarios with a culture (AI feature for meeting speak-up encouragement) or product focus (AI feature about social savings). The exercise followed a structured process: define the scenario, brainstorm and discuss what could go wrong and how to mitigate this along the FORGOOD dimensions.



FORGOOD pre-mortem template available for download soon on the website ("Resources")

- The pre-mortem was well received for surfacing the right issues and good structured space for deep critical thinking and revealing blind spots e.g. to use in retros.
- Interest shown in a "faster" pre-mortem version for firefighting teams limited reflective space (FORGOOD pulse check coming to the website soon).
- Interest shown in tailored pre-mortems for use cases (e.g. employee AI adoption), which are coming to the website soon (function-specific FORGOOD templates, too).
- If you have a scenario you'd like to unpack (anonymised of course) as an exercise in a future workshop, please send it to us in advance and we will build it into the session.

ANY QUESTIONS?

Get in touch with Annabel Gillard | Co-founder FORGOOD

a.c.gillard@lse.ac.uk