OXFORD UNIVERSITY PRESS

Trust & Integrity in Journals Publishing – the role of AI

November 2024

Crisis of Trust

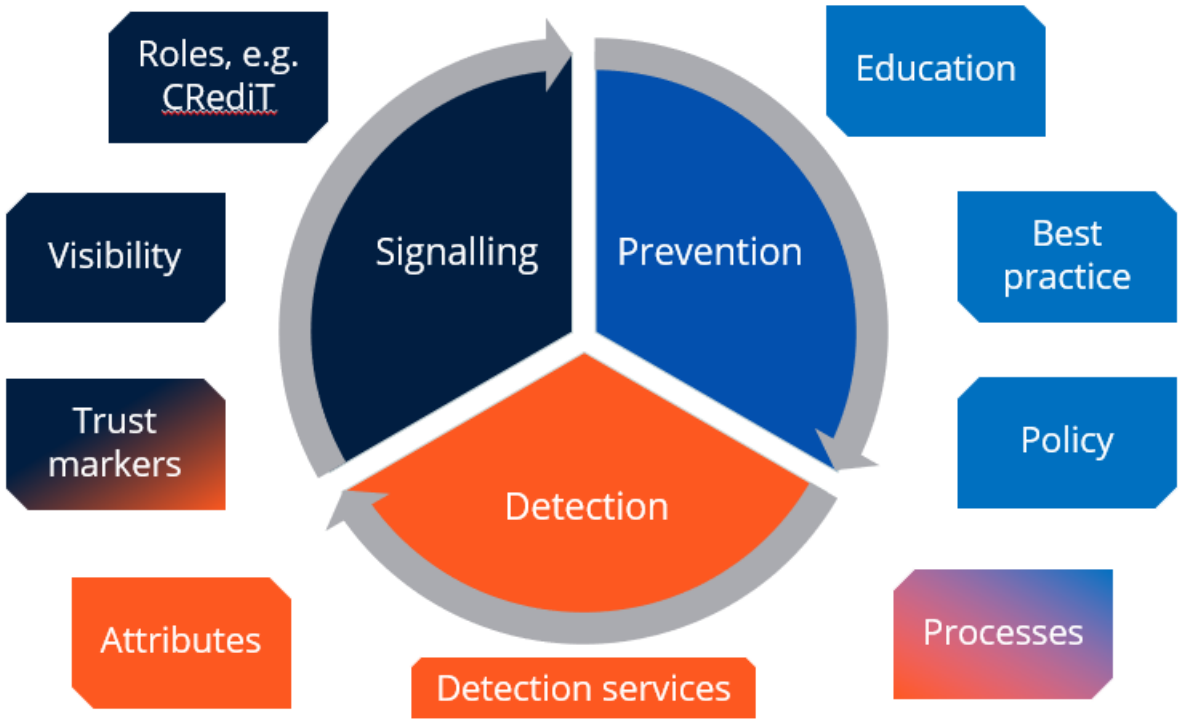
Scale of the problem

- At least 1.5% all published literature estimated to be from Papermills (https://www.nature.com/articles/d41586-023-03464-x)
- Problematic figures in research content estimated at 4% (https://www.nature.com/articles/d41586-024-00372-6)
- Duplicate submissions detection of 1% in STM hub https://scholarlykitchen.sspnet.org/2024/05/23/guest-post-the-research-integrity-hub-connecting-the-dots-in-a-dynamic-landscape/
- >45k retractions in RetractionWatch database

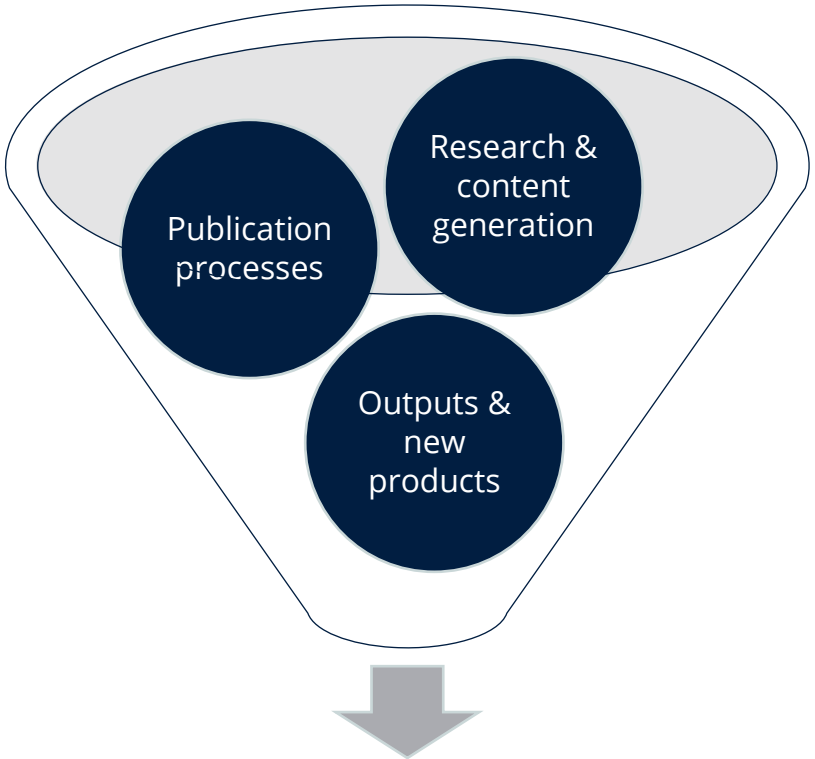
Role of AI

- Technological enabler of problematic behaviour at rapidly increasing frequency
- Risk to integrity of research itself not just publication of work

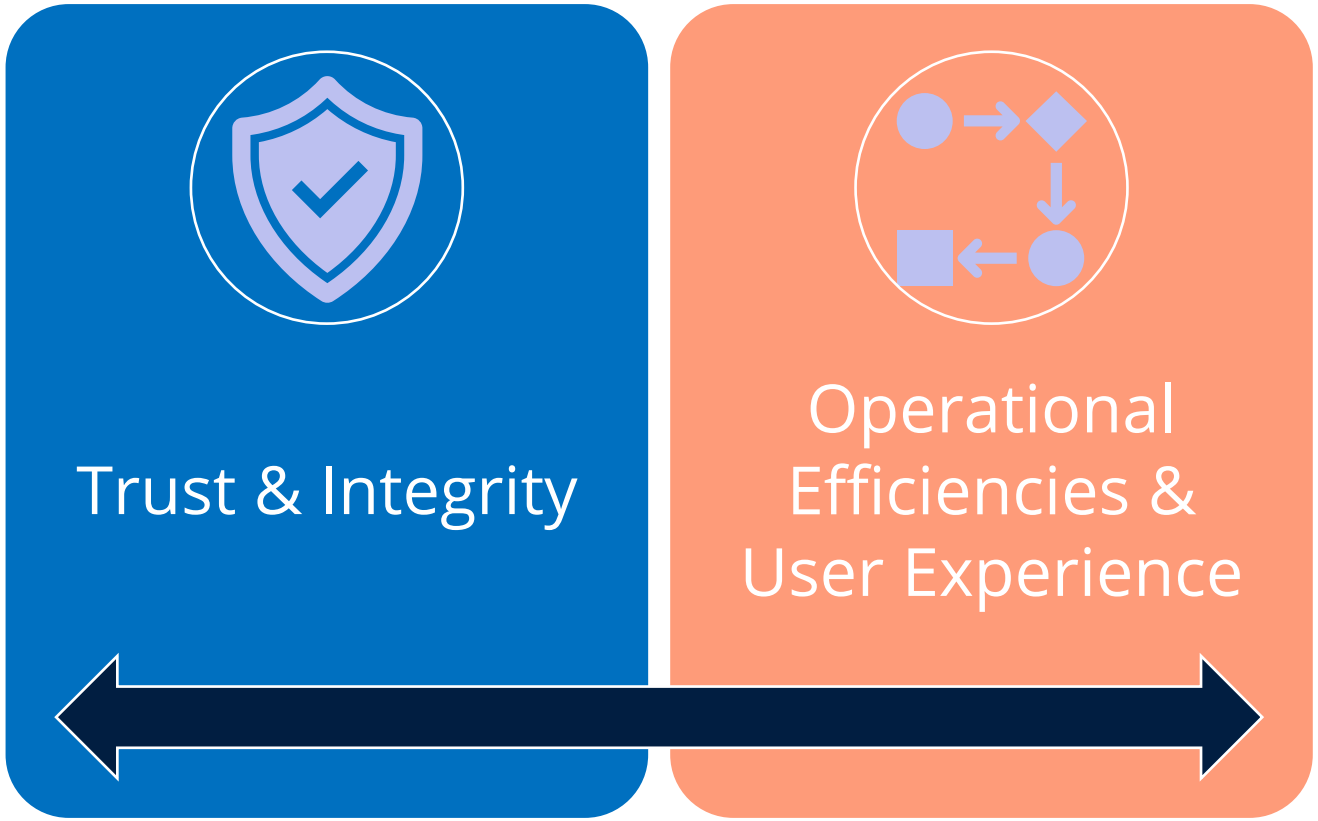
We cannot default to an assumption of trustworthiness



How will AI change Journals Publishing in the near-term?



Unevenly distributed change that will arrive quickly and will bring both opportunities and risks



Quality is critical in AI augmentation

Editorial & Peer Review opportunities using AI

- Integrity review
- Editorial process support
- Peer review support

AI quality controls are critical for effective use

Pretraining

A foundation model is trained with vast amounts of content, enabling it to interpret prompts, develop generalized reasoning, and generate responses in natural human language.

Adaptation & fine-tuning

The model can be further customized or fine-tuned with curated content to optimize performance for particular domains or use cases.

Grounding

Retrieval-augmented generation (RAG) enables an LLM solution to draw on specific sources of content and provide explicit links to the sources used in its response.

AI augmentation to support efficiencies and user experience

31% of post-docs – particularly early-career researchers – already incorporate GenAI tools into workflows (Nature survey, Sep 23)

Image credit: Bing, 2023

Submission

- Authoring support
- Technical checks
- Metadata extraction

Peer review

- Integrity checks & screening
- Reviewer sourcing
- Reporting insights

Production

- Language assessment
- Semantic tagging
- Multimedia generation
- Digital conversion

Post-publication

- Personalized discovery
- Research summaries
- Content recommendations
- Impact dashboards

Exploring human/AI collaborations

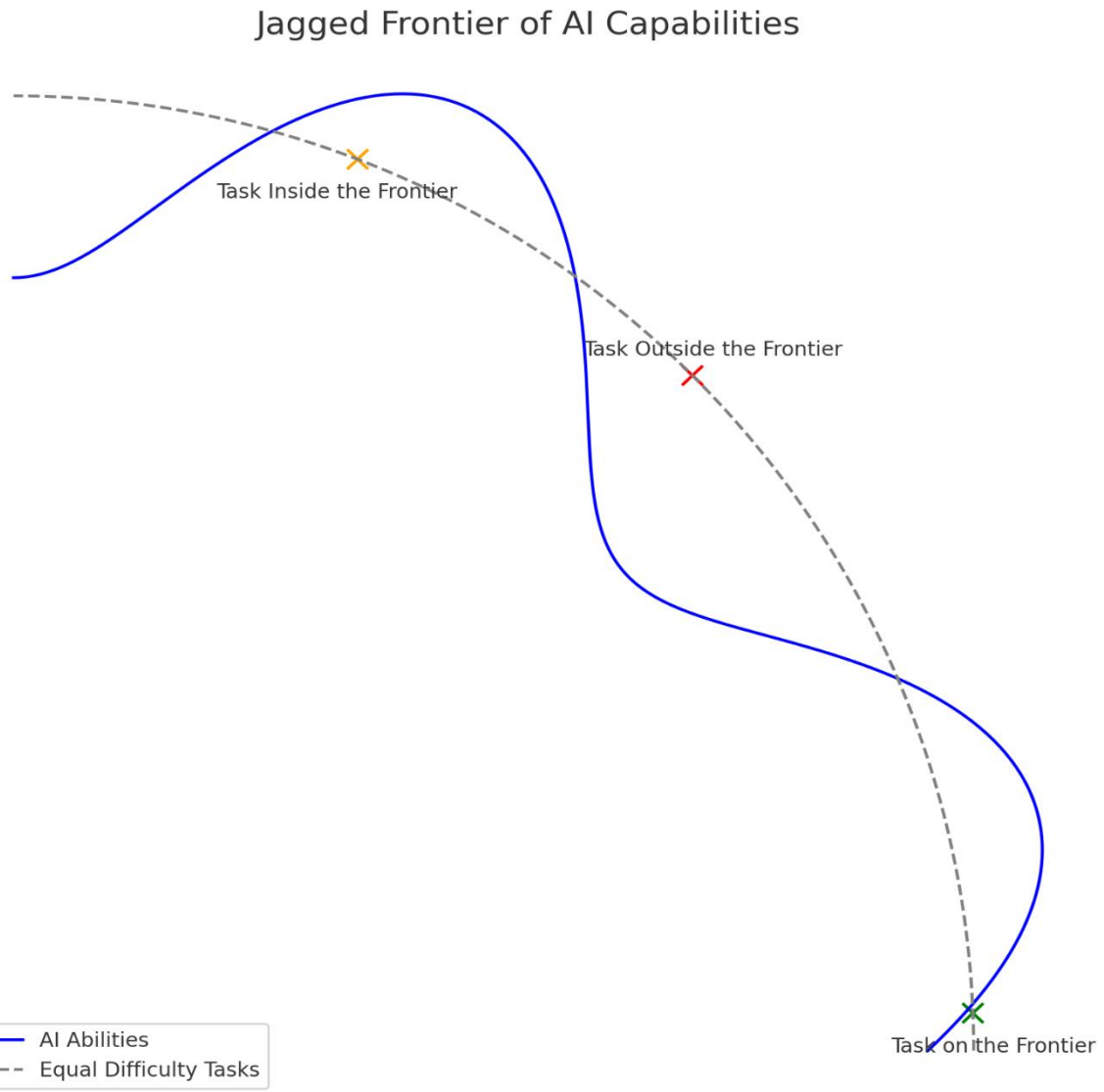
Scholarly publishing is a purpose-driven activity that is facilitated by technology

Incentive structures and academic ecosystems create systematic pressures	Technology can enable activities beyond what humans can do themselves	Human exceptionalism
Technology is a tool used by humans in service of their goals	AI will increasingly behave in ways we don't intuitively understand	AI & technology can help us see and mitigate our weaknesses

A balanced approach to policies

Key considerations for policy setting

- Rapidly changing space
- Policies need regular review
- Policy has a limited ability to influence behaviour
- People's expectations are set by their wider experience across their whole life not just your interactions with them
- Important to avoid excessive duplication
- Policies need to be considered for all relevant stakeholders



https://www.oneusefulthing.org/p/centaurs-and-cyborgs-on-the-jagged

Preserving the value of content through enrichment