

The Good, the Bad and the AI

The New Frontier of Research Management

May 2024
Thomas Gurney, Pure

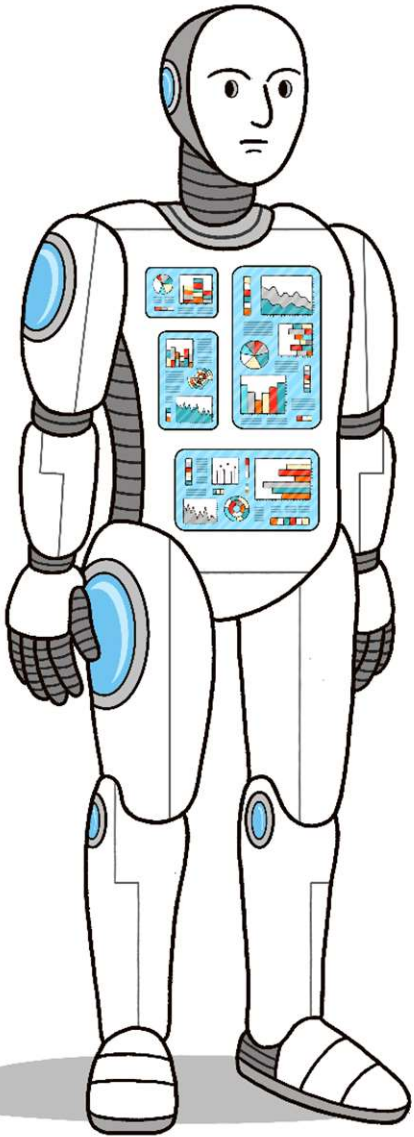




Disclaimer:

This is not a technical, in-depth discussion.

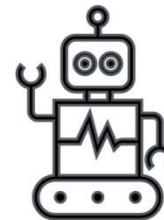
Focus for today



Creation & enrichment
Curation & classification
Process & efficiency



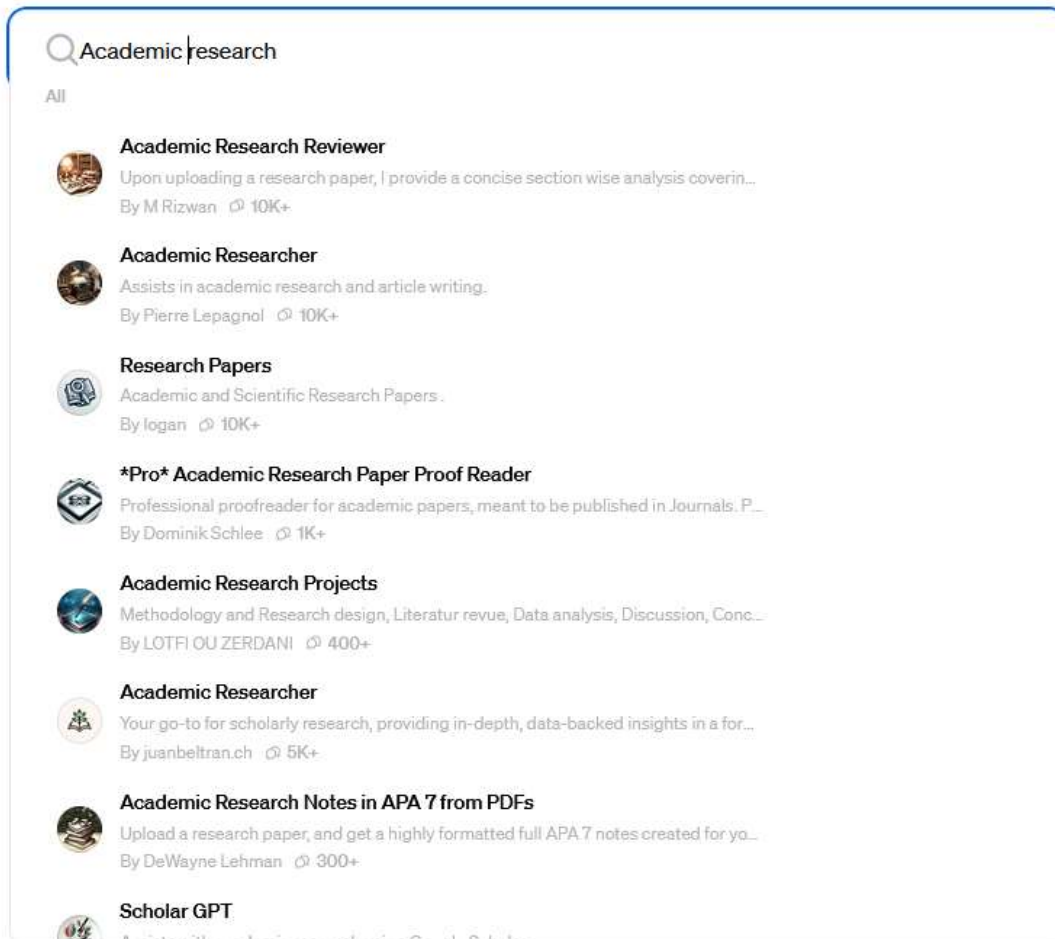
Data integrity & quality
Security & privacy
Over-reliance on automation



Governance
Understanding
Engagement

GPTs

Discover and create custom versions of ChatGPT that combine instructions, extra knowledge, and any combination of skills.



Side note

- Extensive tooling already available for academic research
 - Abstract generators
 - Review analysis
 - Research notes
 - ...

What's our context?

From Code of Good Practice:

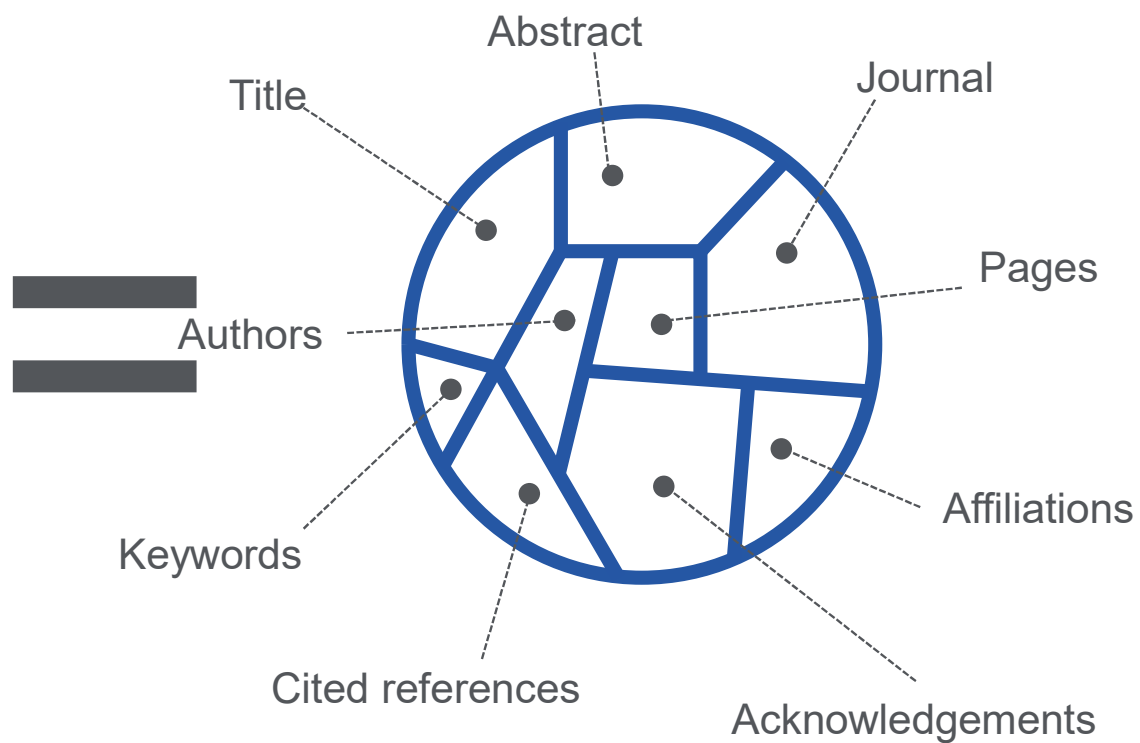
1. Identify potential funding
2. Evaluate research funding
3. Avoid duplication of research activity
4. Establish priorities for research activities
5. Source for scientific publishing
6. Inform educational establishments, industry and the general public
7. Facilitate information access and exchange
8. Facilitate exploitation of research
9. Internal administrative functions
10. Analyse research trends within and across countries
11. Compare institutions within countries in R&D
12. Promote international cooperation
13. Contacts and networks
14. Locate new market for products
15. Locate new products for market
16. Locate persons/organisations with desired skills
17. Administration tool for research projects

Identify
Facilitate
Promote
Analyse
Inform
Evaluate
Exploit
Research





Example: Research output record



Target and source for enrichment

As source

AI ✨

Funding
Collaboration
Impact
Experts

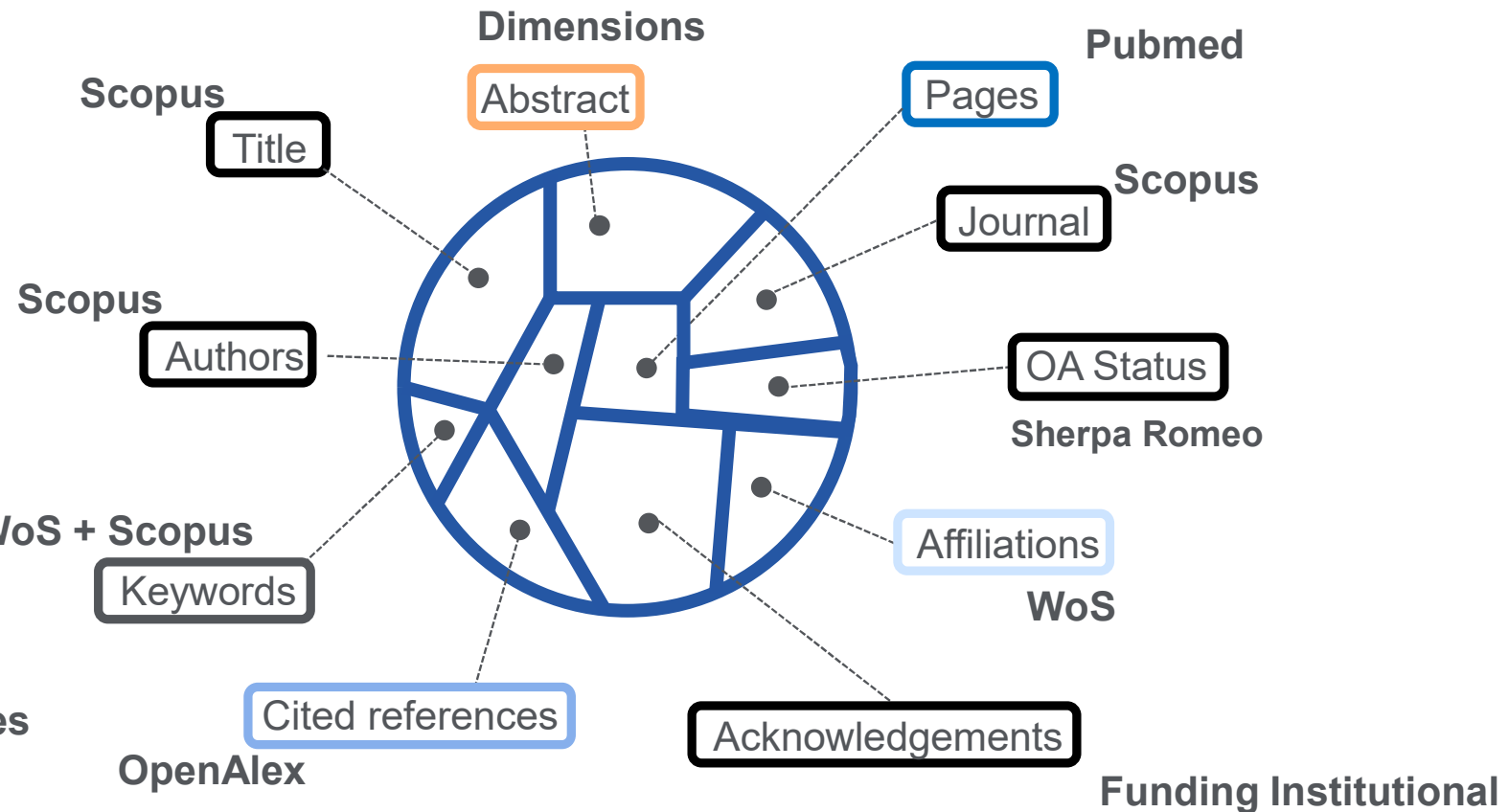
As target

AI ✨

+ WoS + Scopus

Agents:

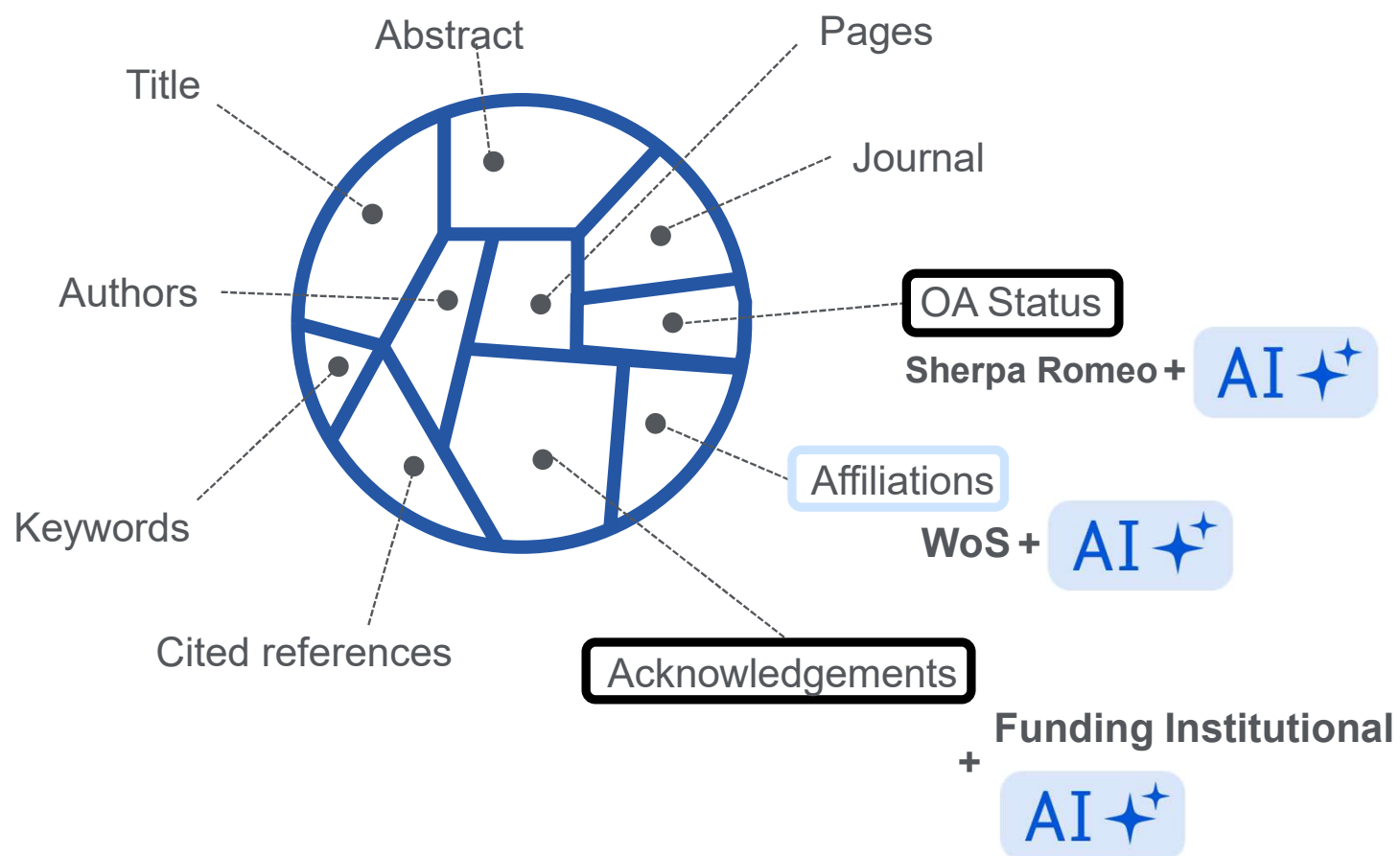
- Funders
- Corporates
- Media



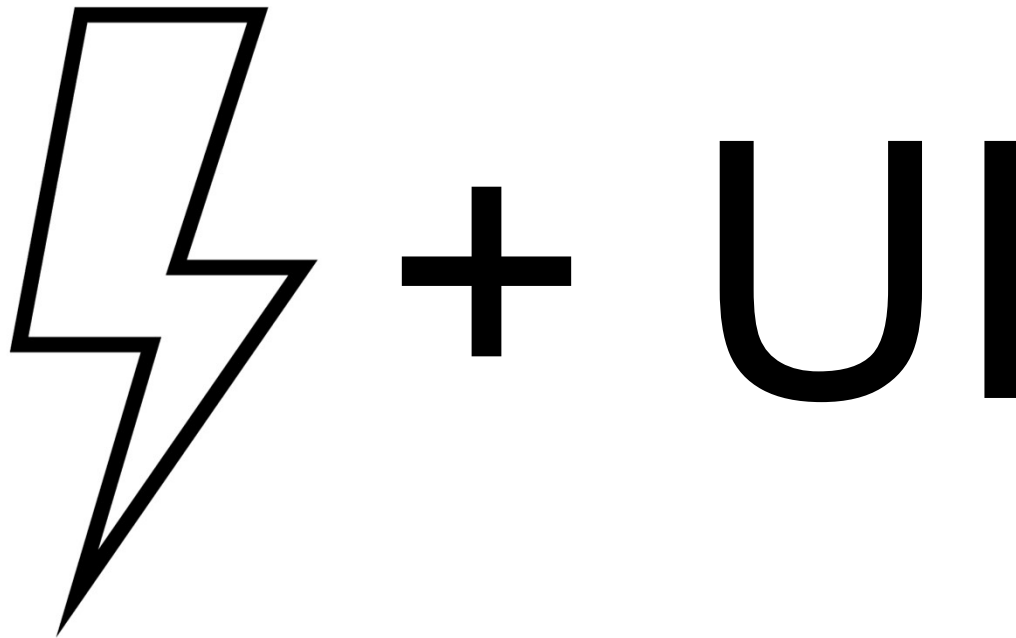
Curation and classification



Cleaning
Compliance



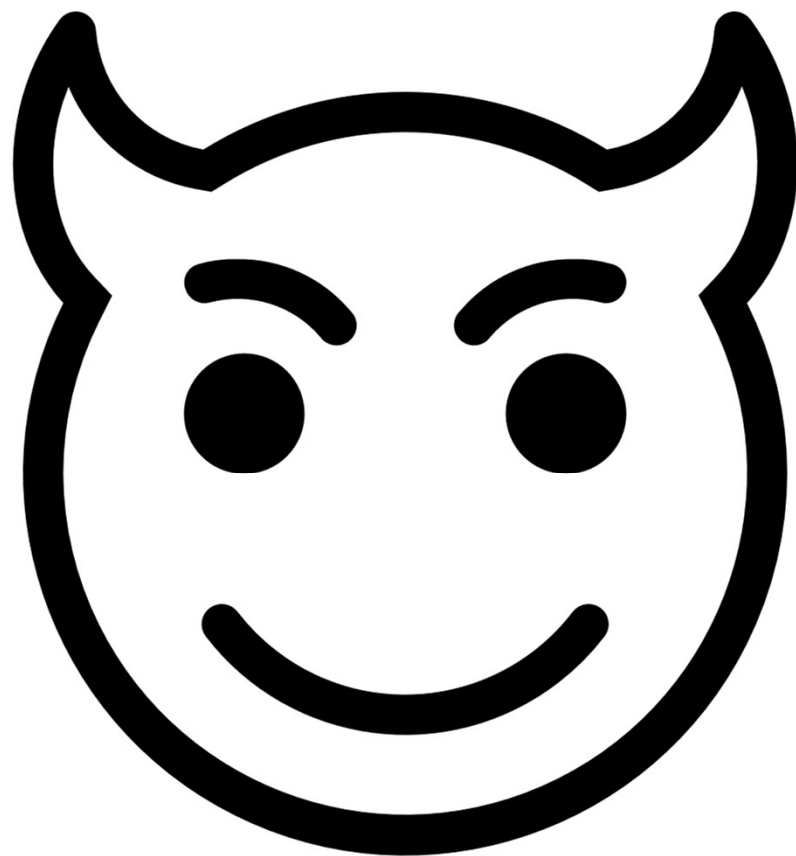
Machine speed



- Incredible efficiency gains are available
- Must be paired with easy-to-use UI components

**Complete
Correct
Connected
Current
Compliant**





Data integrity and quality



“Our biggest source of data quality issues are our users.”

- Anonymous

 Message ChatGPT

ChatGPT can make mistakes. Check important info.



ChatGPT can make mistakes.
Check important info.

Security and privacy



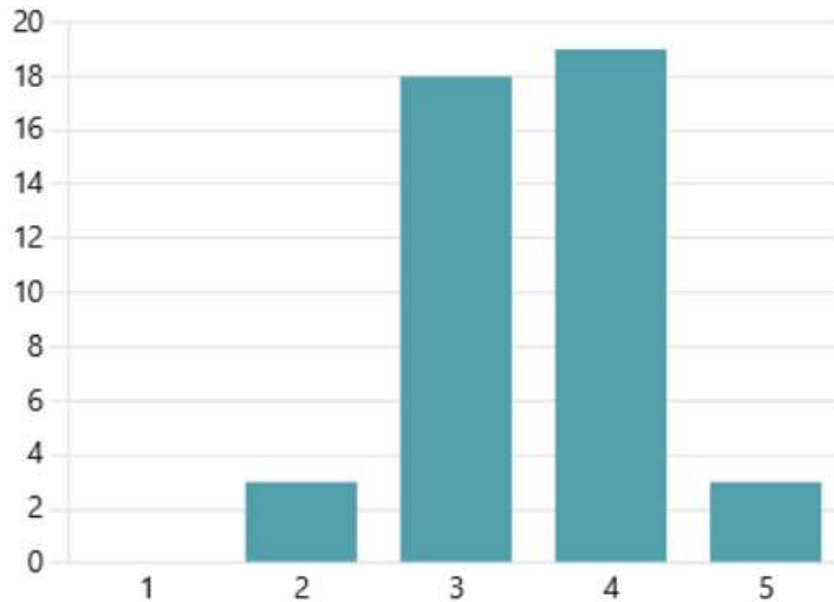
- Personal/Commercially sensitive information
- LLM prompt injection attacks



- How is my data being used?
- How to align GDPR (and others)?



Over reliance on automation

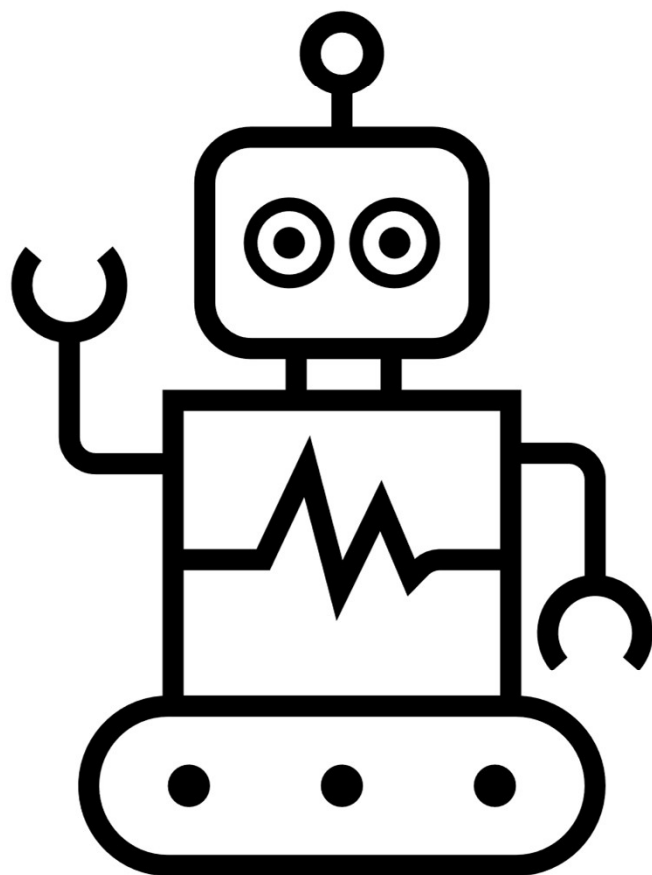


Should data quality be improved by more automation within Pure?

1 - Not at all,

5 - it should be completely automated

- CRIS admins perform the bulk of data management
 - Huge expected increase in content to be checked if GenAI enrichments are implemented
- Human in the loop preferred but resources are scarce



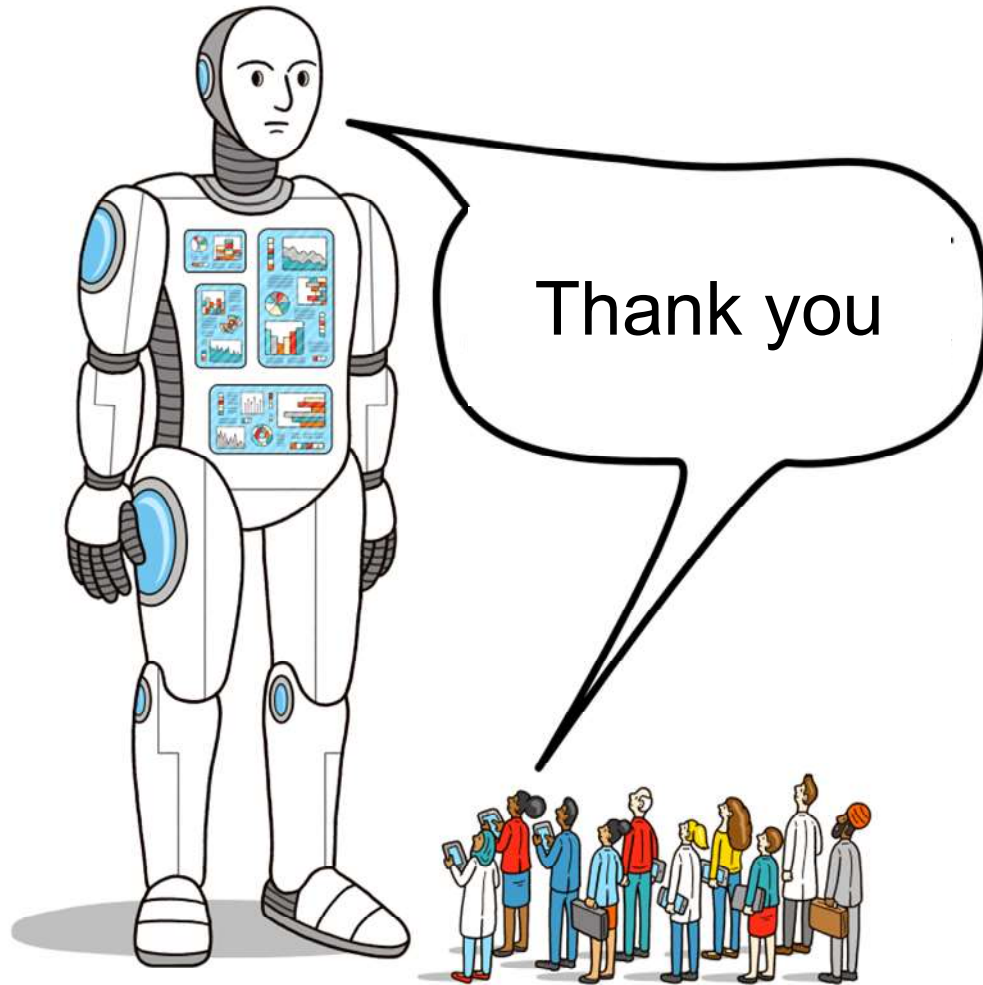
Ready for the future



- Governance
 - Guides for (non)specialists
 - Improved workflows & tagging
- Analysis of security and privacy concerns
 - What data to share where
- Engagement
 - Required oversight



Supercharge | Sabotage



Please provide a snappy thank-you for the end of my presentation to an audience of CRIS users and administrators.



Certainly! For a presentation closing aimed at CRIS users and administrators, you might consider a thank-you that acknowledges both their role and the impact of their work, while also connecting to the theme of your presentation. Here's a snappy thank-you that you could use:

"Thank you all for your time and engagement today. Your dedication not only powers our current research infrastructure but also shapes the future of open science. I appreciate your insights and contributions, and look forward to seeing how together, we can continue to innovate and enhance our research management tools. Let's keep pushing the boundaries of what we can achieve!"

This closing thanks the audience, highlights their importance in the field, and encourages future collaboration, fitting well with the technical and forward-looking tone typical in product management settings.

