AI Policy Editorial Consultation Report.

Compiled by Dr Simon Everett, 16.1.26.

Executive summary.

- **The Open Library of Humanities (OLH) needed to respond to the rise of generative AI use in scholarly research** as other publishers are, as it is no longer an issue that can be avoided due to its ubiquitous nature across online applications. However, we wanted to approach this process as openly as possible, allowing for our editors to contribute to the formation of the policy as a truly community-driven effort in this important area with very conflicting opinions.

- **The policy received one round of internal OLH team consultation, followed by two formal full OLH-wide editorial team consultations throughout 2025**. This gave us a balance of oversight from our own OLH team alongside valuable insight from editorial teams across our journals, which proved instrumental in moving the direction of the policy.

- **Roughly a quarter of our journals responded to the first-round editorial consultation, with around a fifth of our journals contributing to the second round of editorial consultation**. We were grateful for all responses to the draft policy and all feedback was considered extensively, with much added or adjusted in light of editorial comments.

- **The following sections of this document provide our working, including the rationale of policy changes** according to the editorial feedback received, primarily actioned by Dr Simon Everett in the formation of the final iteration of the policy that has been published.

- **The policy will continue to be monitored and adjusted according to additional feedback over the coming months**. This will mean that the OLH is open to any further advice and comments on what would be useful for all parties in the publishing process, which we will consider alongside any major developments in the legalities and ethics surrounding generative AI use.

Editorial consultation round 1.

- **9/34 (26%) of journals responded to the first round of editorial consultation**.

- **Most respondents were receptive of and grateful for the initial guidance OLH proposed**. However, there was a sharp divide between those who have some interest and knowledge about the use of generative AI in academic publishing and those who are leads/scholars in the subject area as it evolves, and therefore have detailed and probing feedback to offer. In particular, this seems to be the case around AI definitions, uses, and declarations.

- **Creating a division between AI tools and generative AI is problematic** because AI tools are considered generative in their operation. Some suggested that the focus should be more on creation of research vs assistance with research using generative AI, as in, what is acceptable and unacceptable use of *generative* AI, with tools as a part of generally accepted writing and research best practice.

- **A blanket declaration statement for any use of an AI tool is problematic because AI tools are now so ubiquitous** it would be difficult and even unrealistic for authors to be able to declare every use of

AI in a formal manner. The suggestions offered for the declaration of generative AI use only and in specific cases of its use that are acceptable chimes with the above point about AI tools being generative in nature.

- **There is an argument to be made for the transparency of the declaration of AI use** particularly for methods and declaration for generative AI use that amounts to more than undertaking minor tasks that also contain margins of error without AI (e.g. delegation of tasks to junior editors or other actors). Asking authors to point to every small use of an AI tool as part of a declaration is not as important as a formal declaration of AI use for research methods that would not be feasible for humans to undertake alone, or for known specific and demonstrative AI output quotations contained within articles.

- **It is becoming increasingly difficult to detect the creation of AI works without using AI to do so**, and even then this is not really adequate presently and may not be accurate or viable long-term. However, safeguarding the research integrity of scholarly work is still important and should be acknowledged. Transparency, awareness of AI's applications, and good editorial oversight by a human actor should be paramount as a first line of defence but cannot remove the problem that AI poses in this respect.

Anonymised points raised (round 1).

- Data classification as a part of gathering datasets using AI models. '"It is acceptable to use AI tools to format the presentation of data. However, the AI tool must not amend, falsify or distort the raw data" (text from OLH AI policy draft 1). One of the key uses of machine learning tools, in digital humanities spaces, is for classification. For instance, a researcher is using 'AI' (machine learning) to identify portions of newspaper re-use and reprints in nineteenth-century US contexts. Likewise, another researcher used a machine learning classification system, trained to perform as well as humans, to identify how much time had passed in a few pages of a novel. This then allowed the scaling of the system to look at thousands of novels to understand whether or not time was accelerating in fiction over a historical span. In both cases, of course, the use of an AI machine-learning system was declared. Indeed, it was specified as part of the method. However, the current OLH policy seems to outlaw this kind of work. Perhaps this could be clarified?'

- AI can be used ethically for more than demonstration purposes. 'In our lab, for example, we commonly use generative bots to review and extract data for us, to run tests, to produce drafts of code, to summarise, and to help compose things such as literature reviews, methodological descriptions, and so on. We don't think that it is appropriate to simply ask AI to design a test for you and then apply it (an instruction "review this document for evidence of Indian English," for example, might be useful as a quick check, but could never substitute for a carefully developed and literature-based census of relevant linguistic features).'

- Use of AI to generate figures/images/art, etc. 'In which instances would it be permitted to publish AI figures, or not? Academics [are] increasingly creating charts using AI and we should address this if possible.'

- The line 'Should a reviewer be suspected of using' should be changed to 'Should a reviewer be found to be using' in the first-round policy draft. 'Being suspected in itself cannot be cause for censure — it has to be concluded that something inappropriate was done. This paragraph overall seems a bit harsh, e.g., "The reviewer must be blacklisted…" There should be some room for error on the part of the reviewer, conversation, reinstatement, etc. The term "blacklist" is not good to use in any case.'

- Determination of AI-generated submissions is becoming impossible to detect. 'As submissions arrive at the office, [...] how are we as editors to determine the provenance of any given text? Again, some portion of it was most certainly machine-written, just like this note. But what if the entire text

was machine-generated? Or substantial pieces of it, if only to add flourish and finish? Surely, we could tell the difference! As it turns out, we sometimes cannot.'

- Mention of tools such as detectors of 'tortured sentences' or 'retracted bibliographical references' not present.

- Changes to declaration statement framing, clarification and positioning. '[The policy] suggests an incredibly narrow (in scope, in time) of AI use. Even today, people are using AI fluidly, constantly, and in very many tasks and steps in the course of the research. There's just no way that you could ask people to truthfully report on every time they use AI. In no time, it would be like asking people to report every time they use Google. (In fact, every time you use Google now you ARE using AI.)'

- Reverse declaration? 'Given that generative AI is NOT allowed, why not just ask everyone to declare that they have NOT used generative AI in the course of the research and writing?'

- Lack of clarity about the need to submit a supporting document with an AI declaration. 'Does this mean that the authors need to submit an additional document? This might not be practical, and authors may easily forget to submit it. Instead, it might be preferable to add a specific field on AI usage to the platform, which could contain the formulations exemplified here. When submitting a paper, besides filling in the information already required, the authors would also fill in the information on AI. An additional option could be added to state the non-usage of AI. Something like "The authors declare they did not use AI of any kind (AI tool, generative AI) in their research."'

- Distinction between 'generative AI' and 'AI tools' doesn't make sense. 'Generative AI is an AI tool. Further, the various acceptable uses of AI tools mentioned (e.g., summarising literature, improving grammar or formatting) are likely to be done with generative AI tools such as GPT, Claude, Gemini, Llama, Midjourney, etc. So nearly all of this is, in fact, about generative AI. The crucial distinction is between unacceptable uses of generative AI - namely, creating research - and acceptable uses of generative AI — namely, assisting the research and writing process in other ways.'

- Concerns about the conflation of generative AI and unethical research practices. 'We aren't convinced by what we see as the conflation of generative AI and unethical research behaviours such as fake authorship, the fabrication of data or citations throughout this draft. The practices that are attributed to AI are practices that existed before commercial chat bots were widely available and they have never been considered ethical. And they are not intrinsic to the use of generative AI. You can fake references without using AI and the use of AI doesn't mean you are faking references.'

- Investigations into AI malpractice. 'In practice most such investigations will be highly inconclusive. AI tools for estimating whether something is AI-generated are themselves highly fallible.'

- Transparency of use. 'Authors should reveal precisely how they are using AI, just as they should how they are using any other tool, resource, or technique. This means citing build numbers and prompts-used, of course, but also declaring when AI has been used in a non-trivial way to help develop substantive arguments and analysis. [...] We think that there are some definite limitations to be required while the privacy rules behind so many applications is unclear. But above, all, we think that the route to accomplishing this is to require thorough and transparent documentation: not to restrict generative AI to certain parts of the workflow or certain tasks, but to require documentation of all non-trivial uses in research, writing, and editing, and a certification that the authors (and editors) responsible for producing and evaluating research using such tools take responsibility for the decisions they have made and have robust controls in place to ensure the results of these tools is properly vetted.'

Changes to our policy draft implemented following round 1.

- **We opted to still pursue a firm but broad and expandable OLH AI policy as clarity is needed for our journals.** Larger (and commercial) publishers have opted for light guidance on good/responsible

use of AI rather than firm policy. However, OLH policy should still be in place that can take aspects of this approach while detailing caution and being respectful of specific generative AI use cases.

- **We removed the division between AI tools and generative AI** and focussed on acceptable and unacceptable uses of generative AI in the broadest possible sense that retained some/most of what we outlined in the first draft.

- **We acknowledged that AI tools are also generative and are ubiquitous** to the point that their use is impossible to restrict given their full-throttle integration in all areas of software and technology, and that their assistive use is also important for accessibility. However, their output(s) should always have oversight by the person using them/reviewing work produced using them to ensure good research integrity.

- **Authors must still claim responsibility for any research methods using generative AI and/or substantial uses of generative AI that form a part of their research** as per legal/ethical guidance from COPE and others.

- **We focused on guidance that fosters transparency of generative AI use** that goes beyond reasonable and minor uses of generative AI tools for research and writing. Minor uses are near impossible to police or even detect, with authors in some cases unaware they are using such tools if they are integrated into software. Public-facing transparency of known, substantial uses of generative AI for research purposes would not be discriminatory in terms of disability/accessibility and would be a matter of methodology and research integrity.

- **We focused on the requirement of declaring substantial uses of generative AI that form part of research methods** such as gathering datasets and classifications otherwise unable or not feasibly able to be done by humans alone, and retain the demonstrative known use cases of generative AI with outputs that are cited as part of an author's argument in their research. If data is proven to be invalid, the author would be subject to editorial investigation and OLH's retraction/removal process should the article be published with this going undetected beforehand.

- **We clarified where, and how, a generative AI declaration statement should be made** by the author, adding that this should be noted on their manuscript at the point of submission, on figures that are created by generative AI for demonstrative purposes such as images or videos (excluding AI formatted tables or charts that have been generated with the assistance of AI from valid data), and also in a submission field in Janeway.

- **We removed the requirement for journals to declare their own use of AI on their sites** outside of OLH's publisher policy as this would be 1) a matter of ethical good practice for editors to ensure they are using assistive AI tools responsibly, and checking all outputs before making final decisions themselves and 2) AI is so seamlessly integrated in some editorial workflows (sometimes without editors knowing, e.g. predictive phrasing, rollout of MS Copilot via Outlook and other services, Google AI searches, etc.) that detailing this could easily become a matter of consternation and confusion with editors who we must trust to oversee their journals and workflows responsibly.

- **We opted to require OLH journals to add to their sites an AI policy statement that they follow OLH's AI policy guidance.** A point will also be added as an item to submission checklists for journals so that authors are aware of the need for AI declaration at the point of submission.

- **We decided to retain the requirement for peer reviewers to provide their reports on the basis of reading an author's submission in full and offering their expert opinion**, and stand by that as a core OLH principle. However, we noted it is unreasonable to expect reviewers not to be using assistive AI tools in the same way authors and editors might, and it is indeed impossible to police this. We decided that it is for editors to evaluate the reports and decide if they might use them or not; develop lists of trusted reviewers and have agency over all decision-making processes based on reports.

- **We created some guidance for our editors to increase their awareness of the use of generative AI in the scholarly publishing workflow.** We acknowledged that such AI tools as tortured sentence and

false bibliographic entry detectors exist and could be helpful for editors if they have the capacity and resources to use them responsibly but not to rely on any one tool exclusively; as noted above, false positives and inaccuracies can occur with using generative AI to detect generative AI. We also added that the use of AI generated images for articles such as thumbnails and cover images is problematic and OLH does not condone the use of AI for such purposes, as rules are in flux surrounding the copyright and attribution of such generated images.

- **We tried to keep the framing of the policy to be as inclusive and sensitive as possible**, as there are many different levels of knowledge on AI throughout our journal teams. Our aim has always been to give our editors confidence that they have some guidance they can trust to follow.

Editorial consultation round 2.

- **6/34 (18%) of journals responded to the second round of editorial consultation.**

- **Although fewer editors responded to round 2**, this was to be expected and the OLH team considered this to be a good sign that we had addressed many of the contentious concerns from round 1. We have also since received some passing feedback from various editorial teams that they are pleased with our changed policy direction.

- **Most respondents were on the whole very grateful that their feedback had been considered for the second draft** and that we clearly worked hard to balance realistic proposals with limited resources and a lack of legal clarity on the use of generative AI in scholarly publishing.

- **Negative responses reflected intense concerns about the rise of generative AI**; concerns about feeding AI models with training material, and how it should be effectively policed were raised. Currently, it is not possible to police the use of generative AI in a practical manner and as a publisher we must wait for further legal and ethical guidance on generative AI use from appropriate sources and organisations before we can put in place firmer guidance.

- **Most of the suggested changes were niche use cases of generative AI that might be better clarified** in the policy, which meant that these could be expanded upon where appropriate by way of examples or clarifications.

- **Some suggested changes reflected more substantial concerns that the humanities as a whole has not yet had to grapple with:** in particular the same sorts of AI analytical methods as social sciences and STEM subjects, which we might consider in more detail to pre-empt the use of generative AI for more substantial use in the humanities for research.

Anonymised points raised (round 2).

- Dictation software for accessibility that uses genAI/LLMs should be listed as an acceptable example of genAI tool use that is assistive technology and need not be declared formally in a statement. Again, oversight of this should be used by the author; however, this is a very 'light touch' use of generative AI comparatively as the author's words are being transcribed verbatim.

- There may also be instances where generative AI can ethically and productively analyse, extend or simulate datasets in a controlled manner. 'In one of the two examples you mention a prompt that would provide the researcher with data on the oeuvre of Dickens. The reason for the use of generative AI here is specifically to evaluate its performance, [i.e. this is a task a human could feasibly handle, albeit impractically]. However, I can also imagine cases where large language models (LLMs) are used to analyse datasets in a controlled environment and with more sophisticated prompts, and

that generative AI as a method in such a case is not part of the research purpose, it is an analytical tool.'

This feeds into the discussion in the policy of research methods and AI declaration needing to align in more complex research cases such as extensions of partial datasets and controlled AI simulations. If so, perhaps it might be useful to include an expansion of this reasoning and an example of this variety as well, to eliminate the potential confusion that only generative AI use which is actually being investigated requires a statement?

- **Generating (synthetic) data to supplement partial datasets does happen**. Here is an example: https://hal.sorbonne-universite.fr/hal-05094577/document. Here, the policy could be that it must be done conscientiously and a statement to the editors should make clear how it was done.

- **Using simulations is also becoming a hot topic** in the humanities. An example is here: https://arxiv.org/abs/2505.19246. The same policy might apply as above.

- OLH asks that substantial and known uses of generative AI should be declared by the author but this should be clearly imparted as being mandatory. At the moment, the phrasing is light on this expectation, so this could be toughened up.

- More substantial discouragement for editors using generative AI to assist with the editorial decision-making processes. There was some concern that we are 'treating editors and reviewers differently' by insisting reviewers read submissions in full but allow for generative AI assistance for editors in the editorial process.

- Add that authors are responsible for errors when using LLMs. Even though this is implicit via all generative AI outputs being the responsibility of the author, this part could be made more explicit.

- Clarification that reviewers using generative AI to summarise articles for peer review is NOT allowed and they should read the submission in full. Current phrasing not as clear as it could be.

- Concern about discouraging generative AI images to be used for article thumbnails, covers, etc. 'How else would it be feasible to easily produce an image that exactly refers to the subject title?' Time constraints are a factor for editors.

- Consider better terminology for 'trivial' and 'non-trivial' AI use. 'There might be a more specific word to be used here to avoid any ambiguity.'

- Using generative AI for translation purposes. There is an issue of translation for those who don't speak English as a first language. 'We are wondering about how we can make our journal more accessible to those from the Global South. With that in mind does there need to be something here about using AI for translation, grammatical sense etc?'

Changes to our policy draft implemented following round 2.

- **Suggested clarifying changes were made** that were reasonable and made sense to incorporate into the existing sections of text.

- **We made allowances for controlled generative AI simulation data and extension of partial datasets** in line with current humanities research and necessities from certain fields, aligning this with the declaration statement section.

- **We expanded the section on AI declaration statements and added a section on complex AI declaration statements** that intersect with research methods, where and how these might be declared, alongside other substantial and known uses of generative AI. This resulted in an expansion and clarification of the more complex uses of generative AI that go beyond the set format(s) already given.

- **We focused on editorial independence** in terms of whether generative AI declaration(s) are satisfactory, aligned with research methods, or made for complex cases of AI use such as simulations, etc.

- **We retained discouraging the use of generative AI images for journal article covers/thumbnails.** OLH's stance is ethically incompatible with generating images in this manner; however, if this is being done by journals, the OLH asks that the image should be marked as AI generated.

- **We took a more firmly discouraging line with editors using generative AI for assistance with editorial decision-making**. However, COPE guidance does not dismiss this outright and the concern raised is that we are treating editors differently to reviewers. OLH's belief is that both editors and reviewers should be undertaking their roles with diligence and professionalism, which means considering an author's submission in full before an editorial or peer review decision is reached. In essence, we have tried to shore this discrepancy up to reflect our publisher expectations on what is required from both editors and reviewers while making allowances for the use of generative AI for reasonable use cases, such as for accessibility purposes.

- **We decided that submitted work translated into English using an LLM should not be required to declare use of generative AI as a mandatory declaration**, as such work will (or should) be subject to the same editorial oversight and scrutiny and peer review process as any other author's work would. Our rationale is that it is no longer feasible or reasonable to be able to detect generative AI use for this specific purpose, and this should ultimately be a matter of editorial judgement for revisions and acceptance. This also reflects the concern that some journals want to attract submissions from places outside of where they usually expect submissions, and where an assistive generative AI tool for a purpose such as translation may be helpful in the interests of promoting such research submissions.