When AI Writes the Complaint: Practical Guidance for Complaint Handlers[1]

Generative AI is changing complaint handling before a complaint even reaches an agency. A person can now ask a chatbot what to do, receive apparently authoritative advice about their rights, be directed to an ombudsman or regulator and generate a polished complaint… all in minutes.

AI assistance can widen access for people who struggle to put a problem into words. It can also flood complaint systems with longer, more legalistic, less reliable and sometimes misdirected material. This report starts from that tension: AI-assisted complaints are not simply better or worse complaints. They are a new operating condition for complaint handlers, requiring practical guidance that preserves access to complaints mechanisms while managing accuracy, proportionality, confidentiality and system capacity.

The paper aims to provide practical and principled guidance for agencies navigating the evolving landscape of Generative AI-assisted complaints. It seeks to address the impact of GenAI tools on complainants and agencies, offering insights into both the opportunities and risks associated with these technologies. The paper offers strategies which will assist the NSW Ombudsman to support agencies in its jurisdiction, while also having broader application in an environment where GenAI-assisted complaints are rapidly becoming normalised.

The starting point is that AI is not simply a tool that assists complaints; it reconfigures the complaint system itself by amplifying volume, reshaping expectations and introducing epistemic risks (hallucination, false authority); while simultaneously improving access and, to some degree, efficiency.

[1] The author would like to thank the several complaint handlers who provided background context at the beginning of this research, William Ruan for this impressive research assistance work in the multi-jurisdictional scoping, Ramona Vijeyarasa and Jenny Ryan for their insightful comments to the final draft, and the NSW Ombudsman Office for committing resources to this important but under-researched topic. Artificial intelligence systems were used for the comparative scan research informing section 2, as described in the accompanying documents. Copy-edit with Harvey (28 June 2026 -Australia-EU version) was tested in the document before the final human review. All text remains the product and responsibility of the main author: Dr. José-Miguel Bello y Villarino, University of Sydney Law School.

This is evident from the very first interaction a potential complainant may have with AI in relation to their case. Instead of searching for help on the internet, citizens may now use AI systems in the form of assistants or chatbots as the first port of call for their queries. The outputs generated by these systems are likely to suggest an agency-specific complaint mechanism and/or the Ombudsman's processes as one of the avenues to address the substance of the complaint, sometimes even offering the email address or a link to the complaint website.

Such an "AI-created prompt" for human action can have systemic effects on the workload of integrity mechanisms open to public requests. For example, the European Ombudsman linked AI systems to a recent increase in complaints received.[2] Closer to us, the number of contacts received by the NSW Ombudsman in 2024–25 (the last reported year) increased by 22%, a 300% acceleration on the previous increase in 2023–24.[3]

Anecdotal experimentation would support this link between generative AI and surges in complaints in NSW. For example, an entry of a test query such as "the NSW Government is unresponsive with my request for funds for X" [X=tested with "solar panels" and "farms"], generated similar replies from the free versions of Copilot, Perplexity, Claude (Sonnet), Gemini and ChatGPT. All of them, after directing to the relevant department/agency channels, suggested escalating to the Ombudsman in their step 3 (4th for Gemini, after going to the local MP and 4th for Claude after suggesting the Energy & Water Ombudsman NSW first). More interestingly, several of the paid interfaces immediately offered the possibility of writing a letter for the user, also reducing the effort required for a contact.

This ease of access comes with obvious trade-offs also observed in NSW. As the 2024-25 Annual Report notes, while overall contacts went up, the share that fell within the NSW Ombudsman jurisdiction declined, largely due to a 50% surge in misdirected complaints.[4] Information received from other jurisdictions confirms this trend.[5] We discuss this issue below, accompanied by ten other aspects of relevance for handling AI-driven complaints

---

[2] 'AI-Driven Complaints Swamp EU Ombudsman' <https://www.lawsociety.ie/gazette/top-stories/2026/april/ai-driven-complaints-swamp-eu-ombudsman/>.

[3] Ombudsman New South Wales, *Annual Report 2024-25* (30 October 2025) 33 <https://cmsassets.ombo.nsw.gov.au/assets/Reports/NSW-Ombudsman-Annual-Report-2024-25.pdf>.

[4] Ibid.

[5] For example the European Ombudsman noted (email communication of 18 June 2026) that they "have seen a 56% surge in the number of complaints submitted and an 86% surge in the number of inadmissible complaints".

(section 3). The purpose of Section 3 is to offer complaint handlers an overview of the potential issues they may face when complainants use AI, but also the aspects that may require guidance from the NSW Ombudsman Office. Most complaint handlers may have experienced one or several of these issues. This document offers the reassurance that those phenomena are real, documented in the scholarly literature and, in some cases, have already been tackled at an institutional level by other agencies.

That analysis is preceded by a short introduction discussing how AI's use can be identified in the complaint handling context (section 1), together with a comparative scan of how other ombudsman-type entities across the world are addressing this issue (section 2).

Finally, we explore the trade-offs facing complainants from more restrictive approaches to using AI (section 4), before finally suggesting a few potential avenues to navigate those trade-offs (section 5), within the existing operational guidance already issued by the NSW Ombudsman.

For the purpose of this report, unless otherwise specified we use "AI systems" and "Generative AI" (GenAI) interchangeably, referring to the digital systems accessible to the general public (with or without payment) capable of producing human-like text. The human-machine interaction can happen directly on the web page, a dedicated app or even be integrated in other existing services (e.g. in Microsoft Word or Gmail). In most cases this interaction takes the form of a "chatbot", with a question/request (prompt) from the potential complainant and an output of structured and intelligible text from the machine. Most commonly the complainant will copy the text into a separate document and then send it through the existing mechanisms (online forms, email, etc.) to the agency to which they want to address the complaint, noting that the approached agency may not be the relevant one.

1 HOW DOES AI-CREATED TEXT APPEAR IN A COMPLAINT?

AI-generated text may enter a complaint process at several different points. In some cases, the complainant may use a generative AI system before lodging the complaint, for example, by asking it to draft, structure or refine their submission. In others, the interaction may be more iterative: a complainant may receive correspondence from the

complaint handler, upload or paraphrase that correspondence into a chatbot, and then use the generated output to prepare a reply. AI may therefore shape not only the initial complaint, but also subsequent exchanges, review requests and escalation pathways. This matters because the complaint handler may be engaging, at least in part, with text that has been mediated through an external system rather than with the complainant's own unassisted account of the problem.

AI Agents tasked by complainants to manage their complaints

AI Agents are complex AI-driven systems that go beyond producing text in response to a single prompt. They are AI-enabled tools that can be given an objective by the complainant, such as "get redress for this complaint", and then plan, draft, send, monitor and revise communications across digital channels with varying degrees of autonomy. Unlike a conventional chatbot, which usually waits for the user to copy and paste its output into an email, form or letter, an agentic system may be connected to email, calendars, document stores, web browsers or complaint portals, allowing it to act on the complainant's behalf within a continuing workflow. This has important implications for the intensity of complaint processes. Where the complainant defines the objective in outcome-based terms, especially as full redress or complete satisfaction, the system may treat persistence, escalation and repetition as rational means of achieving that objective. It may therefore draft the grievance in stronger terms than the complainant would have used unaided, characterise partial or explanatory responses as inadequate, send follow-up emails within minutes, identify several possible complaint-handling bodies and lodge substantially similar complaints with each of them. It may also generate requests for internal review, external review or further escalation whenever the response falls short of the requested remedy, even where the agency has answered the complaint, explained jurisdictional limits or offered a reasonable partial outcome.

The result is not merely a longer complaint, but a more operationally forceful complaint strategy. Each agency response can become an input for the next round of automated drafting, with the agent re-characterising the decision, identifying alleged procedural defects, adding legal or administrative language and producing another polished reply before the complainant has had time to reflect on whether the matter has in fact been addressed.

This may make the whole text and trajectory of the complaint more extreme. Allegations can become more categorical, remedies more absolute, deadlines more compressed and review requests more routine. The concern is not that every AI agent will behave in this way, or that complainants using one are acting improperly. Rather, the risk is that a system designed to achieve a goal may treat compromise, explanation, jurisdictional limits or reasonable closure as obstacles to overcome, instead of ordinary features of administrative complaint handling.

Guidance should avoid assuming a single AI-assisted complainant. The use and effects may look very different for a disadvantaged user relying on a free public AI tool than for a better-resourced complainant using more sophisticated AI support, including AI agents. This piece will not discuss these elements to any relevant detail, but readers are invited to consider at any stage how an AI agent tasked as described before would amplify many of the issues discussed below.

1.1 TRANSPARENCY FROM COMPLAINANT

On occasions complainants may explicitly note that they have used a GenAI system for their complaint and/or leave the chatbot signature (perhaps inadvertently) in the text submitted through the complaint process. For the production of this report, the NSW Ombudsman Office shared with the author a sample of one redacted AI-assisted complaint. This sample explicitly acknowledged the use (Perplexity in this case) over several interactions with the complaint handler. Concretely, the complainant explicitly recognised its use and explained that they had run the reply from the handler through the same tool to understand it better.

In those cases, the acknowledgment may facilitate the processing of the claim as it allows the complaint handler to be aware of the potential shortcoming of that support tool (see section 3 below) and draft the reply taking those points into account (e.g., “the reference to X Act in the AI-generated text is not correct”) mitigating the risk of being perceived as confrontational. In a world where AI assistance tools were normalised and widely accepted complainants would explicitly acknowledge this (e.g. ticking a box in a form) so handlers are aware of this aspect.

1.2 Structural recognisability and formulaic outputs

AI-generated complaints may also be recognisable because they tend to follow relatively standardised drafting conventions. Complaint handlers may notice recurring structures, such as a formal opening, a list of alleged failures, references to statutory obligations, a request for review or remedy and a polished concluding paragraph. The material may also appear disproportionate to the underlying issue: it can be lengthy, highly structured and include sections or arguments that are only loosely connected to the complainant's actual circumstances.

In stylistic and format terms, there may be hints of AI use in English text, such as the use of long bars as section dividers (characteristic of ChatGPT), an excess of Oxford commas or "em" long dashes. A recent review of the literature noted that AI-generated text is "more likely to contain a more formal and impersonal style, signaled by the increased presence of nouns, determiners, and adpositions and the lower reliance on adjectives and adverbs",[6] but these are common characteristics of an "administrative language" style, which could be expected in a complaint.

These features do not prove that AI has been used, and even if AI assistance was not allowed, they should not be treated as a basis for discounting the complaint. They do, however, suggest that AI-assisted complaints may increasingly display a recognisable genre, possibly prompting handlers to assume that AI has been used and potentially, consciously or unconsciously treating the complaint differently.

1.3 Can it be identified with reasonable accuracy through purely technical means?

Even some of the leading systems for detection of AI text show very poor performance, especially in cases where the complainant uses the system to "fix" their text.[7] There is significant work in the field of Computer Science on assessing these tools, especially as there are some environments where AI use is forbidden, such as some education-related uses in assessment contexts and, hence, detection tools are necessary. The accuracy of AI-generated text detection tools varies enormously depending on the

[6] Luka Terčon and Kaja Dobrovoljc, 'Linguistic Characteristics of AI-Generated Text: A Survey' (No arXiv:2510.05136, arXiv, 1 October 2025) 15 <http://arxiv.org/abs/2510.05136> ('Linguistic Characteristics of AI-Generated Text').

[7] Ahmad Hashemi, Wei Shi and Jean-Pierre Corriveau, 'AI-Generated or AI Touch-up? Identifying AI Contribution in Text Data' (2025) 20(4) *International Journal of Data Science and Analytics* 3759, 3768 ('AI-Generated or AI Touch-Up?').

method, the data and the conditions under which they are tested, but, in general, the literature offers no single tool or technique that can be recommended for all scenarios.[8]

1.4 NORMATIVE AMBIGUITY: USEFULNESS VS RISK

In anecdotal conversations about this issue with complaint handlers one sentiment was common: AI-generated complaints are annoying but there is almost always something there. The corollary of this argument is that AI systems could be useful if they were not imposing a higher toll on handlers.

AI use in complaint drafting is, therefore, normatively ambiguous. It can make complaints more difficult to process by adding length, formality and unsupported assertions, but it may also help complainants articulate genuine concerns that would otherwise be difficult to express.

Furthermore, it is impossible to separate the work of the handler from their personal views on AI. A straightforward solution, such as mandatory disclosure of AI use, may have implications on the way the handler "perceives" the value of the complaint. This phenomenon has already been documented in scientific journals and the most recent literature has advocated against mandatory disclosures as they may cause more damage than help in the process of reviewing a submission.[9]

The issue is therefore not whether AI-assisted complaints are good or bad as a category, but how complaint handlers can identify what is useful in them while managing the risks created by AI-generated form, confidence and excess. In the following sections we explore how other complaint handling agencies are dealing with this issue.

2 WHAT WE KNOW ABOUT COMPARATIVE EXPERIENCES

There is only limited publicly available guidance, practice notes or soft law addressing how ombudsman bodies should respond to AI-generated complaints. In late 2025, a Finnish Financial Ombudsman (FFO) surveyed FIN-NET authorities (a network of European national bodies that resolve consumer financial disputes outside court) on this point. Among the 18 responses it received across Europe (including UK) the FFO only identified a few ADR bodies that had *clearly* [emphasis added] received AI-generated

[8] Kathleen C Fraser, Hillary Dawkins and Svetlana Kiritchenko, 'Detecting AI-Generated Text: Factors Influencing Detectability with Current Methods' (2025) 82 *Journal of Artificial Intelligence Research* 2233, 2266 ('Detecting AI-Generated Text').

[9] Mohammad Hosseini et al, 'Disclosing Generative AI Use for Writing Assistance Should Be Voluntary' (2025) 21(4) *Research Ethics* 728.

complaints; and those AI-uses were mainly recognised by stylistic or content inconsistencies. The FFO also reported that these complaints were handled like other submissions, with clarification requests as needed.[10] Furthermore, the FFO noted that no specific guidance was reported for these matters, though members expected the issue to grow with further digitalisation.

Those findings from late last year were confirmed in the systematic scoping we have conducted for this piece. Among 158 Ombudsman-style authorities across Australia, Canada, United Kingdom, Aotearoa-New Zealand, Netherlands, Estonia, South Korea, Denmark, Sweden and France, only one policy document was identified that provided distinct guidance for the handling of AI-written or AI-led complaints, and it was not squarely on point, as it referred to freedom of information requests (FOI) prepared with AI assistance (i.e., not "complaints"). Guidance from the UK Information Commissioner's Office (UK ICO) noted that the growing use of AI is contributing to a rise in the number and complexity of FOI requests. This complexity can extend to follow-up correspondence, increasing the administrative burden on public authorities. [11]

As expected, the UK ICO noted that the use of AI does not in itself affect the legal validity of a request. Public authorities remain obliged to respond to any request that meets the requirements of the Freedom of Information Act, regardless of how it has been produced. Existing FOI principles are considered sufficiently flexible to accommodate these technological developments. However, AI-generated requests may contain inaccuracies, including misstatements of legal rights or references to non-existent case law. The ICO emphasises that such errors do not automatically invalidate a request, but they do require careful handling. Some of the advice it gives is directed at ensuring handlers "understand what the law does and does not say [to] feel confident to push back on inaccurate statements". This requires staff who are appropriately trained and able to identify and address inaccuracies. [12]

More generally, the guidance emphasises that AI-generated FOI requests may require clarification where their wording is unclear or overly complex, and authorities should

---

[10] European Commission - Retail financial Services Unit, 'Minutes: FIN-NET Plenary Meeting 18 November 2025' 2 <https://finance.ec.europa.eu/document/download/f358cada-1789-485c-af6b-4407f7906890_en?filename=251118-fin-net-plenary-minutes_en.pdf>.

[11] UK Office of the Information Commissioner, 'Freedom of Information (FOI) and Artificial Intelligence' (1 April 2026) <https://ico.org.uk/for-organisations/foi/freedom-of-information-foi-and-artificial-intelligence/>.

[12] Ibid.

seek this promptly. Where requests become excessively broad—often due to AI—the cost limit (applicable to FOIs, not often—if ever—to complaints) may be triggered, requiring authorities to provide advice and assistance to help narrow the scope. It also highlights that AI can lead to higher volumes of repetitive, burdensome, or poorly framed requests. In such cases, authorities may consider refusing requests under FOIA provisions where they are excessively burdensome, repetitive, disruptive, or part of coordinated campaigns, including where matters have already been resolved. Finally, it provides a sample text that agencies can place in their website, including the possibility of adding the text verbatim and noting that "this advice comes from the Information Commissioner's Office, the [relevant] regulator" for these requests.[13]

The FOI experience is useful but should not be seen as directly transferable. In the complaint-handling context, the emerging pattern is that AI use is becoming more recognisable, but the ordinary position remains to treat the complaint on its merits and seek clarification where needed.

Other identified materials are also relevant for this discussion, but are not directed at complaint-handlers but at complainants. Notably, these texts are presented as mechanisms to set expectations about how AI should (and should not) be used in drafting submissions. Across jurisdictions—including Australia, the UK, New Zealand, Canada, and Denmark—ombudsman bodies have developed public-facing guidance of this kind. Approaches vary.

Some bodies discourage the use of AI altogether (for example, the New Zealand Privacy Commissioner[14]), while others introduce explicit or implicit limits on length and verbosity (such as the Tasmanian Ombudsman[15] and the UK Independent Office for Police Conduct). Others, including the NSW Ombudsman, adopt a more principles-based approach, emphasising accuracy, confidentiality and reliability.[16]

Relevant parallels also emerge from the private sector. Anecdotal guidance issued to law firms in England and Wales on handling AI-assisted complaints highlights challenges

---

[13] Ibid.

[14] NZ Privacy Commissioner, 'Office of the Privacy Commissioner | Using AI to Contact Us?' (2026) <https://www.privacy.org.nz/tuhono-connect/contact-us/using-ai-to-contact-us/>.

[15] Ombudsman Tasmania, 'Complaints' (18 May 2026) <https://www.ombudsman.tas.gov.au/complaints>.

[16] NSW Ombudsman, 'Using AI Tools for Complaints', *NSW Ombudsman* (2026) <https://www.ombo.nsw.gov.au/complaints/using-ai-tools-for-complaints>.

similar to those identified in other public-sector settings, namely the proliferation of lengthy or complex submissions. That material places emphasis on principles such as proportionality and fairness, which are directly transferable to public-sector complaint-handling frameworks dealing with AI-generated inputs.

3 RELEVANT ISSUES WHICH REQUIRE GUIDANCE

3.1 AI ACCURACY, HALLUCINATION AND LEGAL UNRELIABILITY

A first and central issue is the reliability of AI-generated content. Generative AI systems can produce text that is fluent, confident and legally framed, while nevertheless being inaccurate, incomplete or entirely fabricated. In the complaint-handling context, this may take several forms: the complaint may misdescribe the facts; omit key factual circumstances; refer to legislation that does not apply; misrepresent the relevant legal standard; or, in its most extreme form (the so-called "hallucinations"), cite non-existent provisions, precedents or institutional obligations.

This hallucination phenomenon is most common in generalist AI systems (the usual chatbot), but even generative AI *legal* tools commonly show this shortcoming.[17] The presence of these errors in domain-specific tools may have distinct impacts as complainants may have higher expectations of accuracy (and trust) in a paid legal product.

The implications may go beyond the text being wrong. In some cases, it may appear authoritative enough to require careful checking by the complaint handler. This creates a knowledge (epistemic) burden: handlers must distinguish between genuine legal or factual issues and material generated by a system that can present unsupported claims as if they were established grounds of complaint.

The ICO guidance on FOI and AI explicitly refer to this point and notes that AI-drafted requests "might make references to ICO decision notices or case law that doesn't exist" and warns public authorities to "make sure their information access staff are trained appropriately" so that "practitioners who understand what the law does and does not say will feel confident to push back on inaccurate statements".[18]

---

[17] Varun Magesh et al, 'Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools' (2025) 22(2) *Journal of Empirical Legal Studies* 216 ('Hallucination-Free?').

[18] UK Office of the Information Commissioner (n 11).

3.2 EXPANSION AND AMPLIFICATION OF COMPLAINTS

AI-assisted complaints may also expand the scope, number and complexity of issues that complaint handlers are asked to address. A relatively specific grievance can be transformed into a broader set of allegations, legal claims or institutional failures, sometimes with additional references to legislation, procedural duties or remedies that were not central to the complainant's original concern; or even applicable to the actual grievance.

AI-generated material may also be poorly calibrated to the substance of the grievance. It can add allegations, remedies or legal framing that sit uneasily with what the agency has actually done, making it harder for handlers to assess the complaint proportionately. The task is therefore not only to identify whether the complaint contains a valid issue, but to balance the expanded AI-generated narrative against the actual conduct, evidence and response of the agency.

This form of "claim inflation" does not necessarily mean that the complaint lacks substance. The expanded text may still contain a genuine issue requiring attention. However, it can make the task of identifying the core grievance more difficult, requiring handlers to distinguish between the complainant's underlying problem and the additional claims, categories or legal framing introduced by the AI system.

This phenomenon has already been documented in the private sector. Christensen conducted a series of interviews in an organisational setting and concluded that tools designed to enhance efficiency may instead create inefficiencies by generating excessive low-value content.[19] Generative AI amplified the "amount" of communication (the flow), but often at the expense of clarity, coordination and effective resolution. As a result, critical information (key facts, directives or evidentiary material) risks becoming 'hidden in plain sight', buried within layers of superfluous content that receivers must sift through to identify what is truly salient. In complaint-handling contexts, this dynamic can be particularly pronounced, as increased message volume may obscure rather than clarify case-relevant information.

In a public sector setting, where the law or the expectations of the citizen may impose an obligation to address all elements of a complaint, the main consequence would be an

[19] Emma Christensen, 'From Efficiency to Excess: How Generative AI Reconfigures Organizational Communication' (2026) 46(4) *Controversies of AI Society* 70.

expansion of the time and effort invested in a single complaint without improving the quality of the handling. In a worse-case scenario, it risks that the most relevant elements of the complaint—and the ones that the complainant really cares about—are not addressed to the standard that they would be if the complaint were more specific.

3.3 Increased administrative burden and system strain

AI-assisted complaints may increase the administrative burden on complaint-handling bodies by lengthening processes, multiplying correspondence and creating new bottlenecks. A complaint that might otherwise be resolved through a relatively direct exchange may instead generate repeated AI-assisted replies, review requests or appeal-style submissions (the last point is discussed in more detail below at 3.11). This can extend the time required to resolve a matter and shift staff effort from substantive assessment to managing volume, repetition and procedural escalation.

As before with the text of the initial complaint, reducing the actual required input from the complainant and replacing those with quick (not necessarily short) replies drafted by AI systems may negatively affect the handling of the main issue. First, a reply from a complainant may not include the new relevant information requested, but a regurgitation of the existing text together with whatever new content the system may generate. Second, the concern is not only that individual complaints may take longer, due to the increased back and forth, but that AI use may affect system throughput more generally. If many complainants can produce lengthy, polished and persistent correspondence at low cost, complaint-handling bodies may face congestion at key points in the process.

3.4 False expectations and complainant perception

AI-assisted complaints may also create or reinforce unrealistic expectations about what a complaint process can achieve. General-purpose AI systems often respond in ways that are affirming, confident and solution-oriented. They may frame the complainant's position as stronger than it is, suggest that particular outcomes are likely, or present escalation as a natural next step even where the underlying matter is weak, outside jurisdiction or already adequately addressed.

This phenomenon is linked to the "hallucination" effect noted above. The OAIC guidance on this space explicitly links both when it defines hallucination as occurring

"where an AI model makes up facts to fit a prompt's intent" rather than searching for the most accurate answer.[20] That is, the search for words of the system is likely to prime a reply that satisfies the intention of the prompter rather than contradict or qualify the input received.

This can give complainants a heightened sense that their complaint is not only arguable, but "obviously reasonable". Recent academic research considers that this "sycophantic" effect is built into the AI models in an inextricable way.[21] The implications in complaint handling are broad. Cheng et al. have demonstrated that models' responses were almost 50 per cent more likely than humans to produce sycophantic responses, even where users described unethical, unlawful or harmful conduct.[22] Because users tended to prefer and trust these agreeable responses,[23] AI developers have clear incentives to retain such behaviour despite its risks. Most recent (but non-peer reviewed) academic literature considers that "'delusional spiraling' is an emerging phenomenon where AI chatbot users find themselves dangerously confident in outlandish beliefs after extended chatbot conversations".[24]

Here the problem is not that AI systems may promote "outlandish beliefs" but a sense of entitlement and reassurance that would make it very difficult to see any result of the process besides an absolute acceptance of the AI-induced expectations as a failure. In combination with hallucinations, this sycophantic behaviour may also impose an emotional burden on the complainant, who may approach the process with heightened confidence, frustration or expectation after being led to believe that hallucinated precedent or legal authority gives their complaint a stronger foundation than it has.

---

[20] Office of the Australian Information Commissioner, 'Guidance on Privacy and the Use of Commercially Available AI Products (2025 Update)' fn 18 <https://www.oaic.gov.au/privacy/privacy-guidance-for-organisations-and-government-agencies/guidance-on-privacy-and-the-use-of-commercially-available-ai-products>; citing Fan Yang, Jake Goldenfein and Kathy Nickels, *GenAI Concepts: Technical, Operational and Regulatory Terms and Concepts for Generative Artificial Intelligence (GenAI)* (Report, ARC Centre of Excellence for Automated Decision-Making and Society, 10 July 2024) <https://apo.org.au/node/327400> ('*GenAI Concepts*').

[21] Cody Turner and Nir Eisikovits, 'Programmed to Please: The Moral and Epistemic Harms of AI Sycophancy' (2026) 6(2) *AI and Ethics* 168 ('Programmed to Please').

[22] Myra Cheng et al, 'Sycophantic AI Decreases Prosocial Intentions and Promotes Dependence' (2026) 391(6792) *Science* eaec8352.

[23] Mrinank Sharma et al, 'Towards Understanding Sycophancy in Language Models' (2024) 2024 *International Conference on Learning Representations* 110.

[24] Kartik Chandra et al, 'Sycophantic Chatbots Cause Delusional Spiraling, Even in Ideal Bayesians' (No arXiv:2602.19141, arXiv, 22 February 2026) <http://arxiv.org/abs/2602.19141>.

3.5 TRUST IN AI AND EPISTEMIC AUTHORITY

A further issue is the authority that complainants may attribute to AI-generated outputs. Even where an AI system is only producing probabilistic text, users may treat its response as if it were an informed, neutral or quasi-expert assessment of their situation. This can be especially significant where the output is presented in confident, structured and official-sounding language.

A complainant may therefore take the AI's framing of the issue at face value, relying on it not only to draft the complaint but also to interpret the agency's conduct, assess the strength of their position and decide whether escalation is warranted. The difficulty is that this apparent authority is not matched by accountability. The AI system does not owe duties of accuracy, fairness or candour to the complainant or to the complaint-handling body. It is not independent legal advice, or even professional advice in favour of the complainant, but an output in response to a prompt.

Even if the system then has an extra (or several) layer(s) of iterations and checks for verification of accuracy (which may mitigate hallucinations) and it is prompted to reduce sycophantic effects, it is not likely to produce a balanced approach to the issue, as it will be responding to a one-sided prompt. In that sense, better designed prompts (for example following the suggestions of prompting from the agency as discussed in 5.5 below) may even reinforce the trust in the output.

"Algorithmic authority" or "algority"[25] helps explain why AI outputs can be so persuasive to complainants. The output of the system is coherent, confident and plausible, but its authority comes from the architecture of the system rather than from any direct verification of the complainant's circumstances, the applicable law or the agency's conduct.[26] Verification of the outputs through other AI systems (which are often designed in a similar way and unlikely to give an alternative answer) may even reinforce this.

Recent research on trust in AI helps explain how this authority can consolidate over time.[27] Initial attitudes towards AI may shape whether a complainant is willing to rely on

---

[25] Frida Milella and Federico Cabitza, 'Perceiving AI as an Epistemic Authority or Algority: A User Study on the Human Attribution of Authority to AI' (2026) 8(2) *Machine Learning and Knowledge Extraction* 36 ('Perceiving AI as an Epistemic Authority or Algority').

[26] Arguably AI can "verify" the law by also connecting with any secondary sources and related caselaw on the interpretation of it. What it cannot do is "verify"real-world real-time circumstances and conduct and any relevant digital artefacts that neither the complainant nor the AI has access to

[27] Milella and Cabitza (n 25) 5.2.

a system in the first place, but repeated interaction can deepen that reliance. As users receive fluent, confident and apparently useful responses, they may become more inclined to treat the system as a legitimate source of guidance. This is particularly relevant in complaint handling, where a complainant may use AI not only to draft text but to interpret correspondence, assess whether the agency's response is adequate and decide whether to escalate. The more the task appears suited to AI assistance, the more likely the user may be to accept the system's framing as authoritative.

3.6 Communication and interaction difficulties

AI-mediated complaints may also make communication between complainants and complaint handlers more difficult. Where a complainant relies on AI to draft, interpret or respond to correspondence, the interaction may become less direct and less responsive to the handler's actual questions.

Instead of clarifying the factual issue, providing missing information or narrowing the scope of the complaint, AI-assisted replies may restate earlier allegations, introduce new framing or generate formal but unhelpful text. This can produce a form of communication friction. The handler appears to be communicating with the complainant, but in practice each exchange is filtered through a system that may not understand the procedural purpose of the interaction. This may become apparent when the written communication is replaced with a phone call.

3.7 Loss of human element in complaint processes

AI use may also weaken the human dimension of complaint handling. Complaints are not only vehicles for transmitting information; they often express frustration, vulnerability, confusion or a desire to be heard. From the point of view of the agency, human communication may carry an added level of empathy, difficult to transmit in a written exchange.

When a complainant relies heavily on AI-generated text, that personal dimension may be flattened or displaced by polished, generic language. AI can assist with drafting, but it cannot replace the human experience that gives the complaint its meaning. Some research in the educational domain suggests that systems do not always preserve the tone of the prompt: neutral prompts may elicit disproportionately positive or affirming responses

and negative tones may be amplified.[28] In complaint handling, that matters because tone is not incidental. It can help handlers understand urgency, distress, frustration or uncertainty. If AI systematically moderates or re-colours that tone, it may make the complaint easier to read while also making it less faithful to the complainant's own account.[29]

The Foundation for Financial Services Complaints Institute in the Netherlands "Kifid", stresses this point in its guidance. It notes that they have explicitly decided not to use chatbots: according to the customer surveys that Kifid has carried out "people who call on Kifid attach great importance to personal contact" and as the complaints often relate to major personal and financial situations "direct contact from person to person is often preferred. Personal contact contributes to careful treatment and people appreciate this".[30] From the complainant perspective, it encourages complainants to "do it in your own words" even when using AI, focusing on "what happened, *what are you dissatisfied with*, and how would *you* like to see it resolved?" (emphasis added).[31]

3.8 Platform and variability effects

AI-assisted complaints will not be shaped in the same way across all systems. Although some traits are common (e.g., hallucinations, sycophantic nature) different platforms have different model architectures, training data, safety settings, default tones and product designs. As a result, the same complainant, using the same underlying facts, may receive quite different outputs depending on whether they use ChatGPT, Copilot, Gemini, Claude, Perplexity, Grok, DeepSeek, Qwen or another system.

Some systems may produce short practical advice; others may generate lengthy legalistic submissions, escalation pathways or suggested remedies. Some may be more

---

[28] Li Xiangming and Wei Wei, 'Mapping AI Feedback under Emotional Prompts: A Sentiment-Topic-Network Approach' [2026] *International Journal of Information and Learning Technology* 1, introduction and 5.1 ('Mapping AI Feedback under Emotional Prompts').

[29] The NSW Supreme Court's stance on the use of GenAI in the preparation of evidentiary and expert reports notes that: "Affidavits, witness statements, character references should contain and reflect a person's own knowledge, not AI-generated content" and "GenAI must not be used for the purpose of altering, strengthening or diluting…evidence when expressed in written form". One could consider that it would be possible to substitute "evidence" with "complaint", but the different operating environments and responsibility regimes make the analogy very weak. See Chief Justice New South Wales, 'Supreme Court Practice Note SC GEN 23 Use of Generative Artificial Intelligence (Gen AI)'.

[30] 'Kifid zet AI in als hulpmiddel, mensenwerk blijft uitgangspunt ("Kifid deploys AI as a tool, people work remains starting point")' (12 May 2026) <https://www.kifid.nl/nieuws/kifid-zet-ai-in-als-hulpmiddel-mensenwerk-blijft-uitgangspunt/>.

[31] Ibid.

cautious about uncertainty, while others may present the complainant's position in more confident or affirming terms. In that sense, beyond stylistic traits,[32] the expectations about the process may first be shaped by the model used. If enough complainants disclose the model used, agencies could potentially pre-empt some of the issues linked to a concrete system.[33]

Importantly, as a recent research-based piece in *The Economist* has noted, systems have values embedded into them.[34] The article claims that Chinese-origin generative AI (DeepSeek or Qwen in their analysis) is more deferential to authority, especially if prompted in that language. At a systemic level this may mean that some cohorts of the Australian population more likely to use one model over others would be more or less inclined to put through a complaint, depending on that choice of tool.

This variability also matters for complaint handlers because "AI-assisted complaint" is not a single category. The operational challenges created by AI use will depend partly on the platform used, the version of the model, whether the complainant used a paid or free interface, and the way the prompt was framed. Guidance should therefore avoid assuming that all AI-generated complaints have the same features.

3.9 PROCEDURAL FAIRNESS AND EVALUATION CHALLENGES

AI-generated complaints also raise procedural fairness challenges for decision-makers. Handlers must assess the substance of the complaint without being unduly influenced by the form in which it is presented, which, as noted before, may be difficult, as this may also happen at an unconscious level. For AI-assisted complaints the effects could be felt at both extremes. A polished, confident and legally framed submission may appear stronger than the underlying facts justify, while a visibly or acknowledged AI-assisted complaint may also trigger scepticism about its authenticity or reliability.[35] In both directions, the risk is that the mode of production affects the assessment of merit.

---

[32] See, for example, an analysis of the different traits between DeepSeek and ChatGPT in Hasan MS Jaashan and Wagdi Rashad Ali Bin-Hady, 'Stylometric Analysis of AI-Generated Texts: A Comparative Study of ChatGPT and DeepSeek' (2025) 12(1) *Cogent Arts & Humanities* 2553162.

[33] Models may, however, evolve too quickly and pre-empted guidance could easily become outdated.

[34] 'AI Models' Values Are Very Different from Most People's' *The Economist* <https://www.economist.com/briefing/2026/06/25/ai-models-values-are-very-different-from-most-peoples>.

[35] In a journalistic setting 'labeling a text as AI-written vs. human-written reduced perceived message credibility (d = 0.36)'. Angelica Lermann Henestrosa and Joachim Kimmerle, 'The Effects of Assumed AI vs. Human Authorship on the Perception of a GPT-Generated Text' (2024) 5(3) *Journalism and Media* 1085.

The practical challenge is therefore to evaluate AI-assisted complaints by reference to ordinary administrative justice standards: what happened, what the agency did in response, what outcome is being sought and whether the complaint falls within jurisdiction. The relevant question is not whether the complaint sounds persuasive, but whether it identifies a matter that can and should be addressed through the complaint process. That requires a reflection from the complaint handler to consider any potential biases they may have linked to the AI contribution or the format of the complaint.

3.10 CONFIDENTIALITY

Publicly accessible AI systems raise acute confidentiality risks because any information entered may be retained, reused or exposed beyond the user's control. Once data is inputted, it becomes difficult (or impossible) to track or remove. This is compounded by risks of data leakage, re-identification and adversarial manipulation, whereby models can inadvertently disclose sensitive information or be induced to reveal it.[36]

As noted in the NSW Ombudsman website,[37] complainants should be made aware that AI tools may retain, reuse, or share the information entered into them, creating risks that personal or sensitive data could be accessed beyond the intended context. From the point of view of complaint-handling processes, correspondence from the ombudsman's office may contain confidential material. In this context, AI use by complainants in which that material is uploaded to a generative AI system to better understand it or to prepare a response from the complainant may jeopardise that confidentiality.

Complainants should therefore be warned and supported to exercise caution when using such tools and avoid uploading sensitive information. The NSW Ombudsman website addresses this through an encouragement to contact the responsible officer listed in the correspondence for direct assistance if clarification is needed regarding any aspect of a complaint or the information provided. Although that encouragement may be enough to prevent some instances of violation of confidentiality, properly addressing this issue

[36] Chandrasekhar Uddagiri and Bala Venkateswarlu Isunuri, 'Ethical and Privacy Challenges of Generative AI' in Khalid Raza, Naeem Ahmad and Deepak Singh (eds), *Generative AI: Current Trends and Applications* (Springer Nature, 2024) 219 <https://doi.org/10.1007/978-981-97-8460-8_11>.

[37] NSW Ombudsman, *Using AI Tools for Complaints* (n 16).

for cases where confidentiality matters most (e.g., corrective services, health) may require more compelling guidance.

3.11 Institutional processing and review dynamics

AI-assisted complaints may also affect internal workflows after the initial complaint has been received. In particular, AI can make it easier for complainants to seek review, reframe an outcome as procedurally deficient or provide repeated additional material after a decision has been made. This can shift pressure from the original complaint-handling stage to review, policy or escalation processes, creating bottlenecks where repeated AI-assisted correspondence must be assessed.

The issue is not only volume, but process stability. If each response from the agency can be uploaded into an AI system and returned as a new request, objection or review argument, the complaint may become difficult to close. Handlers may therefore need guidance on when new AI-generated material raises a genuinely new issue and when it simply restates, amplifies or re-packages matters that have already been addressed.

4 TRADE-OFFS

Section 3 set out numerous risks for both the complaints handler and the complainant when AI is used at various points in the complaints process. These risks do not support a simple restrictive response to AI-assisted complaints. The same tools that generate inaccuracy, volume and procedural strain may also help some complainants access the system, organise their account and express concerns they would otherwise struggle to formulate. The policy question is therefore not whether AI use should be accepted or rejected in general, but how complaint-handling bodies can manage its risks without undermining its access-enabling functions.

4.1 Accessibility and support for vulnerable users

Academic literature and government reports have explored in some detail how generative AI and other AI systems could be used to facilitate accessibility to public sector services among vulnerable or diverse users, especially to address language barriers.[38]

[38] For example, a recent study in India addresses this in the context of a multi-lingual, low-literacy environment: Amit Ss Jain, Mahadeo Jaiswal and Bhawana Rathore, 'Bridging the Digital Divide: How Generative AI-Powered Digital Platforms Enhance e-Government Service Delivery' [2026] *Science and Public Policy* scaf097 ('Bridging the Digital Divide'); For a more academic approach to the options offered by AI systems see, Ilaria Mariani et al, 'Improving Public Services Accessibility Through Natural Language Processing: Challenges, Opportunities and Obstacles' in Kohei Arai (ed), *Intelligent Systems and*

However, there is scant analysis of individuals using publicly available models for self-help in environments such as the one discussed here.

An approach to AI use in complaint drafting needs to account for its access-enabling function. For some complainants, AI may help overcome language, literacy, disability-related or cognitive barriers that would otherwise make it difficult to formulate a complaint. This may be particularly important for non-English speakers, people with limited confidence in writing, or complainants who struggle to organise facts into a clear sequence. In those cases, AI is not merely adding volume or complexity; it may be making the complaint process usable in the first place.

4.2 AI AS AN EFFICIENCY-ENHANCING DRAFTING TOOL

In some cases, AI may also reduce processing time where the alternative is not a clear human-written complaint, but a fragmented, unclear or difficult-to-interpret message. In cases involving language barriers, low literacy or poorly organised accounts, AI-assisted drafting may help structure the material, identify the sequence of events and make the requested outcome easier to understand.

For handlers, improved clarity from AI-assistance can make the complaint easier to triage and assess, even if the resulting text still requires caution about accuracy, relevance and proportionality. For example, a complainant who might otherwise send several disconnected messages about delays, poor communication and financial hardship may use AI to consolidate those points into a single chronological account, making it easier for the handler to identify the relevant agency conduct, the impact on the complainant and the outcome being sought.

4.3 A PUSH AGAINST UNNECESSARY ESCALATION LINKED TO PRESSURE TO CLOSE COMPLAINTS

A further trade-off concerns the pressure to close complaints. Complaint handlers cannot be expected to be perfect decision-makers. They operate under time, workload and system-capacity constraints. In that context, AI may sometimes assist complainants by helping them check whether an agency's response has addressed all the points they consider relevant, especially where the reply is technical, lengthy or difficult to

---

Applications (Springer International Publishing, 2023) 272 ('Improving Public Services Accessibility Through Natural Language Processing').

understand. This could help maintain the complaint at the level of the agency where it could be more precisely resolved, without escalating it to other complaint mechanisms, such as the NSW Ombudsman or a thematic-specific ombudsman office.

As noted above, AI-assisted complaints may increase volume, length and repeated correspondence, making closure more important for maintaining system capacity. Pressure to close matters is hence likely to increase. The challenge is therefore to distinguish between complaints that are genuinely exhausted and those that have become procedurally difficult because of the way AI has mediated the complainant's account.

5 POTENTIAL STRATEGIES TO ADDRESS SHORTCOMINGS

5.1 ACKNOWLEDGMENT OF AI USE

Some of the guidance documents discussed in section 2 encourage acknowledgement of AI use without mandating it. At first sight this transparency from the complainant has some practical advantages, the main one being informational. If a complainant discloses that they used AI, the handler may be better placed to interpret the submission, check legal or factual assertions with appropriate caution and respond to hallucinated references without appearing dismissive. It may also help them understand why the text may be lengthy, formal or unusually structured.

At a scale, it could help some well-resourced large-complaint-volume agencies identify patterns across platforms and develop more targeted guidance. However, recalling some of the earlier points in sections 1.1, 1.4 and 3.9, not all the effects of disclosure will be positive. It can facilitate processing, but it can also affect perception. A mandatory acknowledgement risks marking the complaint as less authentic or less reliable, even where AI was used only to overcome language, literacy or organisational barriers. It may also generate a difficult choice for the complainant, because some will not know for sure whether AI has materially shaped their text or, a likely scenario in the future, may have used embedded tools (in text processing software for example) without recognising them as AI. Emotion or a flattening of emotion may be read into the complaint in error. Some complainants may also avoid disclosure if they think it will weaken their case.

The obvious trade-off here is between transparency and stigma. Acknowledgement may be useful as a soft prompt or voluntary support mechanism, particularly if framed

around accuracy, confidentiality and the complainant's own responsibility for the content submitted. However, as a standalone or mandatory solution it may be over inclusive (catches all types of uses even legitimate ones), under inclusive (will not identify those that do not disclose) and procedurally risky (because it introduces biases). It may help handlers manage some AI-specific risks while also reinforcing the very evaluation biases, access concerns and fairness problems identified above. There is still a stigma about AI use in text generation, linked to a "purity myth" of human text.[39] In some environments AI use is still seen with marked suspicion and a complainant may be sensitive to that.

5.2 Explicit Length limitation

The Tasmanian Ombudsman's approach illustrates one way of managing the volume and amplification risks discussed earlier without directly prohibiting AI use. Its complaints guidance imposes length restrictions, requiring submissions to be no more than three pages or 1,000 words, while reserving the option to request further information where necessary.[40] This is presented as a way to support efficient review and timely responses.

Significantly, the length restriction appears alongside specific public guidance on AI use, including a dedicated section titled "Using AI to write your complaint?". Read together, these measures suggest an institutional concern that generative AI may encourage complaints that are longer, more detailed or more expansive than is useful for effective assessment.

The word limit therefore operates not simply as an administrative convenience, but as a practical mechanism for preserving proportionality. It helps keep complaints focused on the key facts, issues and requested outcome, while still allowing the office to seek more information where the substance of the matter requires it. Its advantage is that it addresses AI-related excess indirectly, by regulating the form and manageability of all complaints rather than singling out AI-assisted complainants. Ironically, it could even have the indirect effect on some complainants to push them to AI drafting tools to summarise their claims.

---

[39] Yana Suchikova and Natalia Tsybuliak, 'The Purity Myth: Why Stigmatizing GAI in Academic Writing Is Harmful' (2025) 47(5) *Science Communication* 753 ('The Purity Myth').

[40] Ombudsman Tasmania (n 15).

5.3 ONLINE FORMS

The NSW Ombudsman's online complaint form offers a different way of managing some of the same risks.[41] Rather than imposing an explicit word limit, the form structures the complainant's account through a sequence of focused questions: which entity, what happened, when it happened, why the complainant considers the conduct unfair or wrong, how the agency responded, why the complainant remains dissatisfied and what outcome is being sought. It allows complainants to upload relevant correspondence and agency responses, reducing the need to reproduce large amounts of material in the body of the complaint.

This design may help mitigate some of the risks associated with AI-assisted drafting by directing the complainant towards discrete, decision-relevant information rather than a single extended narrative, although the attachments may cause the same effect. It does not prevent AI use and a complainant may still use AI to answer each field. Nonetheless, it could be expected to reduce the space for unfocused amplification and is likely to make the complaint easier to triage, compare and assess. It could also contribute to the identification of more systemic or related complaints.

Its advantage is therefore similar to a length limit, but more procedural. It regulates the form of the complaint through structured information capture rather than by expressly limiting the number of words submitted.

It is a solution that is easy to integrate in most agencies and could be accompanied by a more direct mechanism (a free-call phone line in the case of the NSW Ombudsman form, always present in the corner of the form). Offering alternative paths is good practice to avoid imposing a way for citizens to interact with the system, which may exclude or be a barrier for some potential complainants.

Setting an online form as a default should be considered against the implicit message it transmits. Extrapolating from academic writing in the context of the welfare state, technical transformations can change the way citizens interact with the public sector, more rarely determined by a letter, a call or a visit to an office, and increasingly by an online interface.[42] It could be seen as a restriction of potential complaints if imposed or

[41] 'Complaint', *NSW Ombudsman* <https://portal.ombo.nsw.gov.au/make-a-complaint-online/>.

[42] Scarlet Wilcock and Georgia van Toorn, 'The Digital Welfare State: Conceptual Limits and Possibilities' (2026) 13(1) *Big Data & Society* 20539517261429193, 2 ('The Digital Welfare State'); Jed

set as a default. Online forms, if combined with low word limits, could be even more explicitly perceived as a barrier.

5.4 CHATBOTS

A chatbot interface can be understood as a more dynamic version of the online form. It can guide the complainant through similar categories of information, what happened, when it happened, which agency was involved, what response has already been received and what outcome is sought; while allowing the interaction to unfold conversationally. In that sense, it may offer some of the same advantages as a structured form, such as focusing the complaint, reducing unfocused narrative and helping gather comparable information. Its main difference is that the structure is delivered through dialogue rather than fixed fields, which may make the process more accessible for some users but also creates additional risks if the chatbot misunderstands the complainant, frames the issue too narrowly or becomes the primary channel through which the complaint is expressed. It does not block AI use by the complainant and, in some cases, even generates a strange situation if the complainant is using an AI agent. They are also becoming more common in the NSW public sector.

A more advanced option would be developing a generative AI interface tailored to complaint handling, that could be reused by different agencies for their needs. It could be a public-sector system designed for this purpose, or an adapted model using a retrieval-augmented generation approach grounded in approved guidance and jurisdictional material. It may look similar to a chatbot from the user's perspective, but it could be better aligned with complaint-handling needs, including jurisdictional triage, confidentiality warnings, proportionality and referral pathways.

The 2024 report *Automated Decision-Making in NSW: Mapping and analysis of the use of ADM systems by state and local governments* identifies chatbots as a recurring form of automation across NSW agencies, internally and public-facing.[43] Although the compendium accompanying the report identified only three chatbots as deployed at the

---

Meers, Simon Halliday and Joe Tomlinson, 'An 'Interface First'Bureaucracy: Interface Design, Universal Credit and the Digital Welfare State' (2025) 59(1) *Social policy & administration* 119, 119.

[43] Kimberlee Weatherall et al, *Automated Decision-Making in New South Wales: Mapping and Analysis of the Use of ADM Systems by State and Local Governments* (Report, ARC Centre of Excellence for Automated Decision-Making and Society, 8 March 2024) <https://apo.org.au/node/325901> ('*Automated Decision-Making in New South Wales*').

time of reporting, it also recorded a larger pipeline of planned or piloted "chatbots for customer service" across state portfolios and local councils.[44]

The report cautions that customer-facing chatbots require careful development and ongoing monitoring, particularly where they are used by vulnerable members of the public. For present purposes, that caution is directly relevant: a chatbot used in complaint handling may structure access and improve consistency, but it may also introduce new risks of misunderstanding, exclusion or over-reliance if it becomes the primary interface between the complainant and the institution.

5.5 Pre-drafted prompts as suggestions for complainants

A further option is for an entity to provide complainants with suggested prompts for use in external AI systems. Rather than simply warning users about accuracy and confidentiality, the agency could offer examples of prompts designed to produce more useful complaint material: for example, prompts asking the system to summarise the facts chronologically, identify the agency responsible, separate facts from opinions, state the outcome sought, or keep the draft within a specified word limit. This may help channel AI use towards the kinds of information that complaint handlers need, while reducing some of the risks of unfocused amplification, legal overstatement or procedural escalation discussed above.

This approach was not identified in the surveyed complaint-handling agencies in section 2 or in a broader web scan. However, governments are already developing prompt guidance for their own staff. For example, the Australian Government's GovAI platform provides Commonwealth public servants with a secure learning environment that includes prompt-engineering modules for tasks such as summarising submissions and drafting responses.[45] This suggests that prompt guidance is institutionally feasible, even if it has not yet been adapted for complainant-facing use.

The limitation is that prompt guidance would not be stable or universal. As noted in section 3.8, different systems may respond differently to the same prompt and the same system may produce different results depending on the model version, safety settings, language of use and whether the complainant is using a free or paid interface. Prompt

[44] NSW Ombudsman, 'Compendium of ADM Systems' <https://cmsassets.ombo.nsw.gov.au/assets/Reports/Compendium-of-ADM-Systems.pdf>.

[45] 'GovAI - Enabling Artificial Intelligence for the APS' <https://www.govai.gov.au/>.

guidance would therefore need to be tested across the main systems likely to be used by complainants, including both free and paid versions, and reviewed regularly as models are updated. Otherwise, there is a risk that official guidance gives complainants a false sense of reliability or produces outputs that no longer match what the agency expected when the prompt was drafted. Furthermore, in a world of increasing systems and models, this task could be extremely resource intensive.

5.6 Use of AI trained to address limitations by complaint handlers

A final strategy is to consider whether complaint handlers should have access to approved AI tools designed to address some of the limitations identified above. If complainants use AI to produce lengthy, legally framed or repetitive material, handlers may need support to identify the core grievance, separate facts from assertions, detect possible hallucinated legal references and compare new correspondence with matters already addressed. An internal tool could assist with summarising material, extracting chronology, flagging jurisdictional issues or identifying repeated claims across correspondence. The advantage is that it may help restore proportionality where AI-assisted complaints impose extreme processing burdens that risk blocking the system. The obvious risks are that handlers may over-rely on the tool, miss nuance or introduce a second layer of AI-generated error into the assessment process.

The most promising aspect of this approach is the design of a tool that could identify hallucinations in relation to published materials cited in the complaints (legislation, policies, precedents, etc.) and flag them in an automated manner, as this is a task where its accuracy could be easily assessed.

This would not remove the need for human judgement. Indeed, it would shift the risk from complainant-facing AI use to institutional AI use. Any such tool would need to operate within existing confidentiality, records management and public-sector AI policies. Complainants would need to be explicitly notified and how the system is used would have to be explained. A more extreme approach in which a system could make recommendations or determine complaint handling does not seem appropriate for this environment.

6 ALIGNMENT WITH EXISTING NSW OMBUDSMAN GUIDANCE

The strategies proposed in section 5 are intended to sit within, rather than displace, the NSW Ombudsman's existing operational guidance. This section maps each strategy against the three most relevant policies, identifying where the proposals are already supported and where they sit in tension with current settings.

6.1 EFFECTIVE COMPLAINT MANAGEMENT GUIDELINES

The complainant-facing strategies align closely with the Effective Complaint Management Guidelines.[46] Structured online forms (5.3), a complaint-handling chatbot (5.4) and suggested prompts (5.5) advance the principle of "information and accessibility" and the objective of making it easy for customers to lodge a complaint, while structured information capture supports early resolution and triage. Voluntary acknowledgement of AI use (5.1) and training handlers to identify hallucinated content fit the Guidelines' emphasis on a skilled and supported team and the duty to assess a complaint on its merits without being unduly influenced by the form in which it is presented.

The Guidelines already contemplate automated written acknowledgements, provided that a complaint handler checks their content and appropriateness, which supports handler-side AI tools (5.6) operating under human oversight. The principal tension is that any length limit (5.2) or default online form must not become an accessibility barrier, given the Guidelines' insistence on multiple communication channels and accommodation of complainants with diverse needs.

6.2 MANAGING UNREASONABLE CONDUCT BY COMPLAINANTS MODEL POLICY

The length and channel strategies find a clear, but qualified, basis in the model policy on unreasonable complainant conduct.[47] That policy expressly authorises restricting overly lengthy written correspondence (for example, a maximum of 15 typed pages) and

[46] NSW Ombudsman, 'Effective Complaint Management Guidelines'.

[47] NSW Ombudsman, 'Managing Unreasonable Conduct by Complainants Model Policy' <https://cmsassets.ombo.nsw.gov.au/assets/Managing-unreasonable-complainant-conduct-a-model-policy-and-procedure.pdf>.

the frequency of contact, providing direct support for managing the AI-driven volume and amplification that the length-limitation strategy (5.2) targets. The alignment is only partial, however.

The important distinction is between restrictions imposed because a particular complainant's conduct has become unreasonable, and general design measures intended to support efficient and effective complaint management for all users. A general length limit, structured form or prompt that encourages concise drafting should not be framed as treating AI-assisted complaints, or lengthy complaints more generally, as unreasonable conduct. Rather, such measures are better understood as ex ante mechanisms for helping complainants present the information needed for assessment in a manageable form. They can support proportionality and timely handling without implying criticism of any complainant or of agencies that adopt reasonable length expectations.

The model policy frames conduct-based restrictions as responses to unreasonable conduct, to be applied "cautiously and sparingly", only in the small number of cases that raise substantial health, safety, resource or equity issues, and only after considering proportionality, the complainant's personal circumstances and the least restrictive option, with rights to warning, review and appeal. Section 5, by contrast, contemplates generalised limits or structured lodgement mechanisms applied at the point of entry to all complainants. For that reason, agencies adopting general limits should ensure that the balance is right: limits should be presented as tools for clarity, triage and timely handling, not as sanctions; they should preserve alternative channels and exceptions where needed; and they should make clear that further information may be requested where the substance of the complaint requires it.

6.3 Generative Artificial Intelligence - Use by NSW Ombudsman officers

This NSW Ombudsman policy[48] governs officers' institutional use of AI, so it bears most directly on handler-facing tools (5.6) and on any complaint-handling interface developed by the NSW Ombudsman (5.4). The safeguards that section 5 attaches to handler tools align with the policy's core controls: accountability for decision-making

[48] NSW Ombudsman, 'Generative Artificial Intelligence - Use by NSW Ombudsman Officers' <https://cmsassets.ombo.nsw.gov.au/assets/Resources/Policies/Generative-Artificial-Intelligence-policy.pdf>.

remains with the officer, all AI-generated content must be fact-checked and cited, and confidential or sensitive information must not be entered into prompts.

The key constraint is that the policy prohibits open-access generative AI from being used to make substantive decisions about complaints, or to communicate with a particular complainant about the handling of their complaint. Any handler tool or NSW Ombudsman-developed interface would therefore need to be a closed-access tool implemented following a risk assessment, rather than an open-access product. The policy is otherwise silent on the complainant-facing strategies (5.1, 5.2, 5.3 and 5.5), so it neither authorises nor prohibits them, leaving a gap that the proposals in this report would help to fill.

7 CONCLUSION

As discussed across this paper, AI-assisted complaints should not be treated as an exceptional or inherently problematic category of complaint, but a central element of a changing operational environment. AI may improve access for complainants who need help structuring or expressing their concerns, but it may generate challenges for complaint handlers to the point of having systemic effects. The main principle of this paper is that guidance must preserve access while supporting clarity, proportionality and efficient complaint management.

Bodies with a regulatory, oversight or guidance role in complaint handling, including the NSW Ombudsman, are well placed to assist agencies in making that adjustment. This could include model public-facing guidance for agency websites and complaint forms, explaining that complainants remain responsible for the accuracy of AI-assisted material, should avoid entering confidential or sensitive information into public AI tools and should focus their complaint on what happened, what response has already been received and what outcome is sought. It could also include internal policy and training for complaint handlers on how to identify hallucinated claims, manage AI-generated amplification and assess complaints on their substance rather than their form.

The key is to keep the balance right. Length expectations, structured forms, suggested prompts and handler-side tools should be framed as general measures for supporting effective complaint management, not as sanctions or as responses to unreasonable complainant conduct. They should preserve alternative channels, allow exceptions and remain subject to human judgement. As AI systems and complainant practices evolve,

complaint-handling guidance will need to evolve with them, with the aim of maintaining a system that is accessible, proportionate and capable of identifying the real grievance within increasingly AI-mediated text.