The trap of using AI as a substitute for 'writing' or 'thinking' (Or why I would rather 'sweat blood' over every word and sentence that I write or choose)
- excerpt from 'Doing a PhD' (2026)

With even basic 'internet search engines' now being upgraded to include some at least indirect AI facility, it is clear that a few basic distinctions are required about 'the appropriate response' to the challenge of 'writing' and 'thinking' in an 'AI-influenced world' – especially when it comes to the example of academic writing and research. Since the transition from the animal kingdom, humans have always in so many various ways found uses for 'new tools' (or related technologies) that are often as much an extension of collective minds as of individual human bodies (e.g. chirographic, print and digital literacies vs. spears, digging sticks, and axes). The functions of 'generative AI' represent an exemplary new dilemma insofar as these would seem (to many people) to be somehow autonomous to and perhaps able to 'replace' the applied intelligence of human thought as well as various other particular activities. We would argue that on closer inspection that even the most advanced generative AI functions at the end of the day still remain (or should remain) 'mere 'tools'. That is, even if often used in confused, mistaken, and even 'self-destructive' or 'counter-productive' ways, 'generative AI technologies' involve related 'non-natural systems' that arguably remain just an extension of human thought and activities (inevitably to some significant degree at least).

As well as in everyday life more widely, in schools and universities around the world 'generative AI apps or agents' have become so widely available that it would perhaps be generally impossible now to just immediately ban their use completely. Much focus is currently being made on the need for a slow-down or moratorium on the development of AI technologies because of concerns about 'rogue AI behavior', the 'loss of control', and related 'security risks'. But in terms of related abuses of AI apps in education, professional credentialism, and 'future work' (or rather the threats to the most effective human learning or critical thinking (and related notions of quality education and knowledge), there seems to be little interest or effort at this stage by either national governments or particular universities to more effectively address the greater challenge of trying to 'draw the line' on appropriate usage, needed safeguards, and related agreed-upon rules or policies. .

It was Socrates who 2500 years ago warned that a general move to 'writing' (chirographic literacy) was dangerous because it threatened the 'dialogic' capacity of humans to think, to remember, and to communicate with others directly in focused relevant ways. This is ironic because Socrates' prescription to this historical challenge (by asking or focusing on relevant questions and problems to construct productive 'threads of inquiry') perhaps remains just as relevant in 'the AI age'. This is especially so in all education, and also in more specialised or advanced human 'knowledge-building' (exemplified by the process of 'doing a PhD'). It was Socrates' who best exemplified (and himself practiced) the kind of critical thinking skills and deep learning practices which many of the more aware educators are now warning must be retained to provide similar antidotes to already emerging general abuses of AI in education, knowledge more widely, and related 'accreditation' processes (in the wider society as well as in schools and universities).

In practical terms then, how might we start to try and make relevant distinctions between an appropriate or legitimate use of AI as 'a tool' and emergent tendencies for superficial, threatening, and counter-productive 'misuses of AI' that are starting to seriously undermine our education systems and even our societies more widely? I think we need to begin with what many

agree is the needed first step of better recognising from the start that such counter-productive abuses of academic standards and practices (such as credentialism and 'publication as mainly just a numbers game' as well as descriptive and ad hoc tendencies of 'research-based writing' generally) were there long before the particular dangers of 'generative AI' came on the scene – to just further exacerbate earlier tendencies already in place. Put another way, long before AI came along to be used in similar fashion, many postgraduate researchers (and learners generally) would more typically undertake a 'research process' (especially 'a search of the literature') that started and proceeded (from the general to the particular) with descriptive 'topics' - rather than 'questions or problems' transforming that process with some relevant design or plan.

In relation to 'research-based academic writing' (and 'knowledge-building') more specifically (and 'education' more widely), I think we need to start by acknowledging the following: that generative AI apps and agents may indeed be useful (and potentially legitimate and constructive) tools for helping to speed up and assisting the accessing and harnessing of relevant background 'data and information' for relevant purposes. And after this concession, we can begin to tackle the more pivotal dilemma and challenge of coming up with sufficient safeguards and adequate standards to distinguish '*the use vs. abuse of AI'*. This central dilemma perhaps lies in how, as is generally well-known already (and causing school and university and other educational administrators great anxiety) some particular AI apps and agents are being promoted as 'being able to write a book or even a PhD' (and not just short school assignments or even a regular 'work reports') – as well as in professional work also now regularly replacing copyrighters, accountants, and legal assistants (to name a few of those increasingly affected).

With the 'doing of a PhD', the 'line is crossed' when generative AI functions are not just used as 'helpful background tools' but effectively become 'co-creators' (if not the main 'creator') of a written text such as a dissertation (as well as other related 'digital products'). We can see this with a platform like *Grammarly* which encompasses both a general 'chatbot' to assist and also 'specialised agents' that guide particular parts or stages of the process. Grammarly also includes such dedicated agents as AI-*proofreaders, paraphrasers, plagiarism checkers, citation finders*, and *data analysts* - whilst other similar platforms and apps likewise often include dedicated 'AI writer' chatbots or agents which can also function as *academic editors, structural outliners, literature mappers*, and so on**.** In other words, such functions have been especially designed to conceal apparent plagiarism or what we would normally in the past call 'cheating'.

It was not long ago in universities that academic staff typically used a program called *Turnitin* in to see where students had simply copied from other published research articles (especially in introduction and literature review sections). In addition to other tools to avoid such detection, generative AI editors are now particularly good at modifying text sections instantaneously for different purposes and audiences (i.e. especially including the key skill of 'paraphrasing' that would always catch out cheating or plagiarising students). Although it has not until now been a regular practice in universities, some others (as well as myself) have always distinguished between '*surface plagiarism'* (the copying of some particular texts or information as 'word segments) and '*deep plagiarism*' (the copying or reproduction of central and/or supporting ideas). Whilst AI renders the detection of surface plagiarism effectively redundant, it cannot help a PhD student if they are unable to demonstrate their own original and effective inquiry design along the lines that can be optimised using the emergent method. And therefore, we have to hand an ultimately reliable and available basis on which to also start 'drawing the line' – and work back from there to outline further related protocols and practices to protect and encourage 'academic integrity' in the 'doing of a PhD'.

The trap of using AI as a substitute for 'writing' or 'thinking'

The situation is of course a little more complicated than it seems. Ultimately there can and will be better programs and apps in the future able to adequately tell whether some writer (and speaker also) is using AI tools or technologies to some degree. And even if we allow some reasonable and appropriate background use of 'helpful AI tools' (as I think we will need to), then this will be indicated with the appropriate related tools also. This is when we will also have to return to the question of 'where do you draw the line' - e.g. in terms of levels, degrees or percentages of 'use' (or perhaps that should be 'mis-use'?

Personally, I think it would be more productive to come at this matter from a different direction. I think the real issue is one of not just the deep learning capacity to '*be able to put ideas in your own words'* (as Richard Feynman also once put it), but really of establishing and then further protecting or otherwise maintaining how this might be linked directly with what I would call '*personal or individual voice'* in writing and also thinking as well as in speaking[1]. Societies and/or particular institutions will need to act soon to starting coming up with appropriate protocols (and related methods of detection) to both encourage and recognise that a student or scholar is able to (or has been able to) maintain a consistent, natural, and original 'voice' – one that is experientially and/or conceptually grounded in the deep learning process by which they have been able to transform mere information (or even 'opinion') into constructive and transferable knowledge.

I acknowledge that such needed protocols and standards will also need to allow some *'reasonable use of AI tools'* (especially for certain types like students for whom 'English is a second or foreign language' having to produce in English). And so, I would not demand that students or scholars be simply banned from doing so. However, there is a reason perhaps worth thinking about why I personally do not normally directly or knowingly use generative AI technologies to 're-write' any of my thoughts or ideas in words. It links to a related reason why I would also generally recommend to others that they try to avoid doing so where or to what extent reasonable or even possible. This is that a general antidote (to 'the AI challenge') lies I think in encouraging new generations of learners, scholars and even 'citizens' everywhere to achieve '*deep learning insights, solutions and outcomes'* (i.e. 'optimal knowledge-building') through the regular effort of 'putting ideas into their own words'.

When I write anything that is important or significant (especially when dealing with complex issues or sensitive topics), then I strive to '*sweat blood'* (i.e. concentrating) over the selection and combination of words and ideas into the most relevant focus and cohesive extended effort that I can in the situation. Time of course is often a pressing constraint – especially in terms of the need for further editing and revision to end up with a final draft that I am happy with. This usually applies to every sentence and paragraph as well as longer sections, chapters, articles or even books. Whilst I know that certain AI apps and agents might assist with coming up with something more apparently impressive and immediately 'polished', the regular use of such tools will I think tend to dilute, to sanitise, and to generally undermine my efforts to be more precise, relevant, and meaningful (and also 'polished' if sufficient time). In other words, at 'a certain point' both in particular pieces and also generally speaking, I know that I would perhaps lose a significant degree of the capacity of '*my voice'* (i.e. the ability to put ideas into 'one's own words') to inform

[1] As further discussed in especially 5.15 below, I think Socratic thinkers like Paul Ricoeur and Hans-George Gadamer would understand this as some variation of the universal 'dialogic voice' function. This is partly reflected in the quote by Socrates on page 111 below (footnote #63). Cf. also our 2027 book Richards, C. *Dialogic Rhetoric: Interpreting the (AI proof) interplay of metaphor, narrative models, and cultural contexts of human language-use.*

the various stages of the process of '*producing knowledge'*.[2] And so, the related challenge of *'where to draw the line'* will need to somehow protect and encourage this capacity in every individual, whilst also allowing the reasonable use of 'tools' to assist..

The main danger of *'writing using AI tools'* is not just that this will tend to be vague, descriptive, formulaic, shallow, and somewhat predictable (as many know already). As suggested also by the footnote below, the real problem lies with how the 'drafting' process can be effectively eliminated with the '*AI polish'* of instantaneous grammar correction and parenthetic revisions to order for different purposes or audiences. Put another way, those who circumvent the various stages of thought as well as writing involved in producing a PhD will be especially missing *the 'distillation process' (the 'productive struggle'*) which is really needed *to transform mere 'data and information' into effective knowledge grounded in relevant understanding as a 'deep learning' process*. Insightful commentators are aware that the greatest problem of 'AI reliance' is that too many people then make of assuming '*that access to information is the same as having knowledge or understanding'*. A version of this is what is known as *'the google effect'* (and also 'd*igital amnesia*') – the psychological tendency to forget relevant information easily obtained. This can be a dangerous situation to be in as it tends to lead to *the kind of 'close-minded ignorance' that Socrates warned against 2500 years ago*.

[2] A basic tenet of the '*process writing model'* (discussed in *'Doing a PhD'* and also '*Words, Ideas, and Optimal Knowledge-Building*) once promoted in certain schools and even universities is that learners learn better as they 'write' – that, after appropriate preparation through prior reading and reflection, they might attempt a 'stream of consciousness' first draft of a piece of writing. According to early exponents of this model like Donald Murray and James Britton (who tended to focus on creative or imaginative rather than expository or more 'conceptual' modes of writing), they should do so without worrying about correct spelling or grammar - as this can provide a basis for further drafts that, through relevant editing and revision efforts improve upon as well as correct the initial draft. Later variations of the 'process writing model' used in higher education (e.g. the Flower & Hayes model, 1981) emphasized the writing process as a series of planned drafting stages (including particular functions of editing and revision) towards completion or publication. Our own emergent method outlined below adapts the process writing insight. But it does so as a semi-structured process in terms of an effective 'thread of inquiry' design organized around a relevant pivotal question or problem.