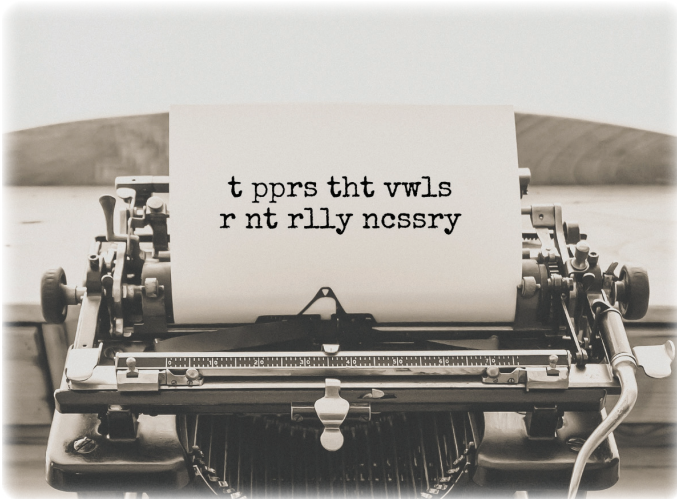


Editorial

t pprs tht vwls r nt rllly nessry

That may well be the least intelligible way that I have ever begun an editorial. Alternatively, perhaps after a little pause, in your mind the letters above became translated to ‘it appears that vowels are not really necessary’. For anyone used to text-based conversations conducted on mobile phones, this is hardly a revelation, although a wider range of abbreviations is used in that context (e.g. c for see, 2moz for tomorrow).

Credit: ha11ok, pixabay via Canva



The opportunity to omit vowels offers a simple way to compress a message. E is the most commonly used letter in the English language, accounting for over 10% of all letters. Precise values depend on the corpus, but each of A, I and O comfortably account for over 5% of letters. U is something of an outlier amongst vowels, accounting for roughly 3% of letters. Nevertheless, simply removing all vowels would easily save at least 25% from the length of a message.

This removal of the most common letters is consistent with Shannon’s information theory. Oversimplifying somewhat, this theory indicates that less frequently used symbols (or letters) contain proportionately more information content. Obviously, there are much more sophisticated compression schemes than simply omitting vowels, and information theory helps us understand their properties as well as theoretical limits.

This understanding is highly valuable, across a wide variety of applications. However, I have recently been thinking about other types of situations where an explicit (or sometimes implicit) desire for optimisation might not be appropriate, because it removes opportunities for individual expressions of richness and diversity. I suspect many such situations involve interactions with people. In these situations, inappropriate optimisation may unintentionally exclude communities who are already marginalised.

For example, during the early days of the World Wide Web many sites did not provide alt-text for images, thus preventing screen readers from providing a full collection of information to visually impaired users. Thankfully, with the advent of accessibility standards, this type of omission is largely a thing of the past. Inclusion of alt-text is relatively simple, and visually impaired users are a relatively easy community to identify. Neither of those observations is intended to minimise the harm

that can be caused when information is not provided and the community is marginalised. However, the observations are intended to prompt us to consider whether newly developing technologies are excluding groups of people who might be hard to identify and doing so in ways that are more difficult to pinpoint.

There are, I believe, indications that tools included in Microsoft Word (and other similar programs) have changed the way that people write [1]. Over time, this has led to a more standardised way of writing. In some situations, standards can be very helpful, if not lifesaving: connecting wires with incompatible voltages is almost never a good idea!

There is also a strong argument that the suggestions provided by Microsoft Word, covering both spelling and grammar, have increased productivity and widened access to opportunities, both of which are positive outcomes. However, there is also an argument that these benefits have been achieved at the cost of making writing more homogenised and less diverse. This may have the effect of making certain forms of expression less common and, in some sense, less valid. Communities that are already marginalised may become even more disadvantaged as a result.

To be clear, even though I’m using Microsoft Word as an example here, it is far from an isolated example; I worry this may be happening more frequently than many people imagine.

These considerations become especially important, I think, in the context of Artificial Intelligence (AI) and, especially, Large Language Models (LLMs). Over recent years, the power of these models has grown at an incredible rate. Their ability to write text, or create images, is now commonplace. But, in developing LLMs, a decision has been made that one form of text is better, or otherwise more valid, than another form. This could easily disadvantage minority communities, whose voices are atypical and, according to one interpretation of information theory, more interesting because of it.

I need to admit that my thoughts in this matter have been strongly influenced by a paper co-authored by Timnit Gebru, who was once co-lead of Google’s Ethical AI Team, and Émile P Torres [2]. For me, this was one of the most thought-provoking papers I have read in quite a while. In the spirit of full disclosure, I should also note that I reached the paper via reading *The*

Future by Naomi Alderman (Fourth Estate, 2023), which was also thought-provoking (and entertaining).

I fear I’m poorly qualified to summarise Gebru and Torres’s paper, so I’m not even going to make an attempt. If you’re at all interested, then I’d encourage

you to seek out a copy yourself. I will, however, share how my thoughts have developed, with the caveat that they are still developing, and that people may well disagree with them!

Firstly, I think we need to carefully consider inputs. In the context of LLMs, there has been significant discussion about prompt engineering, as if this is somehow similar to learning the precise syntax of a programming language. Given the general way interactions with LLMs are intended to occur, for example, using largely natural language, I’m not sure this analogy is appropriate. More generally, there may be a risk that, rather than building AI that is as clever as human beings, we are dumbing ourselves down to a level current AI can handle, requiring standardised forms of writing and expression that may further marginalise already disadvantaged communities.

... inappropriate optimisation
may unintentionally exclude
communities ...

Secondly, I think we also need to carefully consider outputs. In developing LLMs, AI practitioners are, explicitly or implicitly, making judgements that one form of expression is, in some way, better than another. Unless we are very careful, unintentional bias may drive outputs to represent the most common expressions, leading to further marginalisation of under-represented communities.

If they are correct (which, as noted earlier, they may not be), then these thoughts could be representative of a small part of an important problem. That problem will demand an interdisciplinary response, an important part of which is likely to fall to mathematics. More specifically, I think our community could contribute in (at least) three critical ways.

1 Value and measure diversity: We already have an appreciation of the value of diversity, for example, via the use of ensemble models. Our techniques for measuring diversity should provide a mechanism for quantifying the variety and richness of both inputs and outputs.

2 Optimise to make clear choices: Our understanding of optima should allow us to understand how much variety and richness is being traded away, either in terms of standardising inputs or promoting particular types of output. This should allow decision approaches that reduce diversity to be made as clear choices, so that these are then only adopted when the benefits significantly outweigh the disbenefits.

3 Automate to retain richness and diversity: We have developed a range of techniques for working with large datasets and for creating synthetic data. These should provide the means to automate processes that maintain richness and diversity, throughout the development process and in the way that users interact with LLMs.

Given the potential growth, and widespread utility, of LLMs, these contributions (and others) could be of critical importance. If we fail to make them, I worry that everything, and everyone, will be poorer as a result.

Rob Ashmore CMath CSci FIMA
Defence Science and Technology Laboratory

REFERENCES

- 1 Woollaston, V. (2023) The surprisingly subtle ways Microsoft Word has changed how we use language, *BBC Futures*.
- 2 Gebru, T. and Torres, E.P. (2024) The TESCREAL bundle: Eugenics and the promise of utopia through artificial general intelligence, *First Monday*, vol. 29, no. 4.

Crown Copyright © 2025 Dstl. This information is licensed under the Open Government Licence v3. The views and opinions expressed herein are those of the author and do not necessarily reflect those of the Defence Science and Technology Laboratory.